THE AMERICAN JOURNAL OF PSYCHOLOGY

Vol. LXXIII
MEZES HALL, THE UNIVERSITY OF TEXAS
AUSTIN, TEXAS

TABLE OF CONTENTS

ARTICLES AND NOTES

CONTENTS

Book Reviews

(The reviewer's name appears in parentheses after the title of the work.)

Agostino Gemelli

(see page 155)

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. Stanley Hall

Vol. LXXIII MARCH, 1960 No. 1

AN EDITORIAL ANNOUNCEMENT

It is with pleasure that we announce the assocation with this Journal, of Professor Leo Postman, University of California at Berkeley, as coeditor; and of Professor T. Arthur Ryan, Cornell University, as editor of the Book Review Department.

Twice before, from 1911 to 1919 and from 1926 to 1939, the editing of the Journal has been a quadripartite affair. The first Board of four co-editors was provincial. It was centered in New England, as psychology itself was pretty well centered at that time. The second Board marked the break, since at least half of the co-editors were located west of the Hudson. The present group, as now composed, extends over the entire country. One of our members is in New England, one in the East, one in the Mid-Southwest, and one on the Pacific Coast, the broadest sectional representation that the Journal has as yet enjoyed.

Professor Ryan succeeds Professor George L. Kreezer, who has at two periods, 1942-1944 and 1957-1959, served the Journal well as editor of its Review Department. Our thanks go to him as our welcome is extended Professor Ryan.

M. E. Bitterman
E. B. Newman
Karl M. Dallenbach

ON VISUAL MASKING (METACONTRAST): DICHOPTIC OBSERVATION

By Paul A. Kolers and Burton S. Rosner
V.A. Hospital, West Haven, Connecticut

If two concentric visual forms are presented in sequence for brief supra-threshold durations, the presence of the first form is not reported at certain durations of the pause between them. This effect has been called visual masking.[1] If, instead of concentric black forms, successive, neighboring flashes of light are presented, the brightness of the first flash appears markedly reduced. This latter, related, effect has been called metacontrast.[2] Several ambiguities exist regarding the empirical characteristics of these effects, and very little is known about the parameters affecting visual masking. This paper describes the experimental resolution of some of these limitations.

Historical. In his study of masking, Werner presented a small disk tachistoscopically and followed it with a concentric ring whose inner border coincided with the border of the disk. The arrangement which Werner found optimal for an apparently 'retroactive' disappearance of the disk was a repeated cycle of presentations at the following range of durations:[3]

Disk:	12– 25	m.sec.
Pause:	120–240	m.sec.
Ring:	12– 25	m.sec.
Pause:	280–560	m.sec.

With this temporal sequence Werner found that the disk was not seen; but an analogous effect with the ring did not occur. Reversing the order of presentation of the figures, that the ring came first, permitted both figures to be seen.[4]

* Received for publication April 21, 1958.

[1] Heinz Werner, Studies on contour: I. Qualitative analyses, this Journal, 47, 1935, 40-64; P. G. Cheatham, Visual perceptual latency as a function of stimulus brightness and contour shape, *J. exp. Psychol.*, 43, 1952, 369-380; H. H. Toch, The perceptual elaboration of a stroboscopic presentation, this Journal, 69, 1956, 345-358.

[2] Mathew Alpern, Metacontrast: Historical introduction, *Amer. J. Optom.*, 29, 1952, 631-646.

[3] The durations given represent the range of optimum conditions taken across several *S*s, not the range of durations at which disappearance occurred for a single *S*.

[4] Descriptions of related effects have appeared in the literature for many years. Dodge demonstrated 50 yr. ago that lines placed in the pre- and post-exposure field of his tachistoscope inhibited the perception of a word presented on the same part of the viewing screen. Raymond Dodge, An experimental study of visual fixation, *Psychol. Rev. Monog. Supp.* 8, 1907 (No. 35). Exner previously had found that the

In interpreting the results of his 30-odd experiments, Werner assumed that formation of a figure has a time-course and a spatial gradient, and that the latter is highest at the border of the figure. Then he argued that the disk would have just begun forming at its border when the ring appeared; that is, that the disk required from 132 to 265 m.sec. (the sum of its duration plus the pause between disk and ring) to form. When a contiguous ring was presented early enough after a disk, the ring used the partially formed contour of the disk in forming its own contour, which left the disk contour-less and therefore invisible. When, however, the ring is presented first, both forms remain visible because "the ring, in this case, is already in the first stage of development, which permits the contour of the disk to be built up as a separate configuration. Therefore, the whole disk can be seen."[5] This analysis implies that the ring has a developmental advantage over the disk because the former has two borders, and therefore two gradients of formation which summate, while the disk has only one border and therefore only one gradient of formation.

Metacontrast is the converse of masking: where the latter consists of the apparent lightening of the first of two black figures, the former consists of the apparent darkening of the first of two flashes of light. One might question whether 'black' is due only to the removal of light and whether, therefore, inhibition of a black form on a white ground can be regarded as analogous to inhibition of a flash of light on a dark ground.[6] In answer, Werner has shown that the masking effect does also occur with white figures on a black ground.[7] Therefore, in this paper we shall regard the two effects as complementary.

It is helpful at this point to summarize the empirical characteristics of these phenomena about which there is uncertainty. This may be done under the following four headings.

(1) Repetition, and therefore the rate of repetition, seems to play an important role in the effects Werner obtained. From this it might be inferred that any ring acts to inhibit the subsequent rather than the previous disk, and that the mechanism is one of a proactive 'contour inhibition' rather than of a retroactive 'contour appropriation.' One kind of proactive effect has already been named paracontrast — the inhibitory action a figure exerts on a succeeding one.[8] The existence of this effect, which was apparently unknown to Werner, would lead one also to expect an inhibitory action of a disk on a ring in the disk–ring sequence. The question remains how much of Werner's findings depend on this proactive effect.

(2) In Werner's interpretation, absolute contiguity of the forms is assumed to be a requirement for the inhibitory process to occur,[9] but Alpern and Fry reported the existence of an inhibitory effect over angular separations at least as great as 1°

---

brightness of a white figure on a black ground depended upon the length of a pause after which the figure was followed by a uniform white field. Sigmund Exner, Ueber die zu einer Gesichtswahrnehmung nöthige Zeit, *Sitzb. Wien. Akad. Wissen.*, 58, 1868, 601-632. Other variations, and the history of the effects, have been reviewed by Alpern, Cheatham, and Toch.

[5] Werner, *op. cit.*, 43.

[6] Cf. L. M. Hurvich and Dorothea Jameson, An opponent-process theory of color vision, *Psychol. Rev.*, 64, 1957, 384-404.

[7] Werner, *op. cit.*, 41.

[8] Alpern, *loc. cit.*

[9] Werner, *op. cit.*, 43.

of visual angle between the test and inhibiting contours.[10] These latter authors, however, used flashes of light to the dark-adapted eye, while Werner presented black forms to, presumably, a light-adapted eye. Since Werner has already shown that the masking effect occurs with white forms on a black ground as well as with the reverse, it would be surprising if the masking effect with black figures should occur only when the figures are contiguous.

(3) The mode of viewing required for experiencing the various effects is of considerable theoretical interest. One may distinguish at least three modes of viewing: monocular, binocular, and a third which we shall call "dichoptic."[11] In the first of these both stimuli are presented to the same eye; in the second, both stimuli may be presented each to both eyes simultaneously; in the third, one stimulus is presented to one eye only and the second stimulus to the other eye only. In one form of dichoptic observation 'corresponding retinal points' can be stimulated to achieve a fused image, if two identical disks are used as stimulus-figures. Changing one of the disks to a ring would then make the remaining disk and the ring appear concentric. The uncertainty of experimental results regarding mode of viewing concerns only the case of dichoptic observation as just described. Alpern reported that he was unable to obtain any metacontrast effects with dichoptic observation; but Werner with a stereoscope and Toch with appropriately crossed Polaroids have reported finding dichoptic inhibitory effects.[12]

(4) Alpern reports that metacontrast cannot be obtained when the stimuli are made to fall on the fovea; apparently comparable results were obtained by Toch.[13] Werner, however, reports that the masking effect can be obtained in the fovea though not as readily as it can in the periphery. Since Werner reports only the absolute sizes of his stimuli, but not their angular subtense, it is possible that his stimuli were large enough to extend into perifoveal regions and be inhibited there. Crawford, however, has found that a 0.5° test-patch fixated foveally can be inhibited by a 12° conditioning-patch that follows the former up to temporal separations at least as great as 100 m.sec., and Cheatham has reported analogous results.[14]

The experiments reported here were undertaken first to establish whether dichoptic masking does in fact occur. Success in reproducing this phenomenon then led us to study quantitatively some of the temporal

---

[10] Alpern, Metacontrast, *J. opt. Soc. Amer.*, 43, 1953, 648-657; G. A. Fry, Depression of the activity aroused by a flash of light by applying a second flash immediately afterwards to adjacent areas of the retina, *Amer. J. Physiol.*, 108, 1934, 701-707.

[11] Consistency of usage would require that the first two be called 'monoptic' and 'dioptic' respectively.

[12] Alpern notes in his Historical Introduction that other investigators have reported dichoptic effects, notably Stigler, and Baumgardt and Segall. The latter reportedly found, as Werner did, that a different time-course marked the two modes. Stigler's data appear not to have been published. See Alpern, *op. cit.*, 1952, 644. See also Werner, Studies on contour: Strobostereoscopic phenomena, this JOURNAL, 53, 1940, 418-422; and Toch, *op. cit.*, 350-352.

[13] Alpern, *op. cit.* 1953, 654-656; Toch, *op. cit.*, 349.

[14] B. H. Crawford, Visual adaptation in relation to brief conditioning stimuli, *Proc. Roy. Soc. Lond.*, 134B, 1947, 283-302; P. G. Cheatham, Visual perceptual latency as a function of stimulus brightness and contour shape, *J. exp. Psychol.*, 43, 1952, 369-380.

and geometrical conditions for its occurrence. These conditions were: (1) the durations of the stimuli and of the pauses between them; (2) the angular separation between the stimuli; (3) the shape of the stimuli and their order of presentation; and (4) their retinal location.

METHOD

Apparatus. A four-channel electronic timing circuit was used to control the duration and intensity of the light-sources illuminating the stimuli.[15] These sources were mercury-argon gas-discharge tubes coated with a magnesium-tungstate phosphor. The optimal latency and decay of these lamps to full on and full off, as judged from the oscilloscope trace produced by a 929 phototube at the surface of the lamp, were 1.4 m.sec. each. The light from the lamps, measured with this phototube and oscilloscope, was a square wave of constant amplitude for durations as short as 2 m.sec.[16] The precision of the timer was ±3% for the durations used. The entire system was operated from a Sola regulating transformer.

Let the four visual fields used be numbered 1, 2, 3, 4. The first stimulus to one eye was illuminated by two tubes in Field 1, and the second to the other eye by two tubes in Field 3. Field 2 was used for the inter-stimulus interval (ISI), that is, for the pause between the first stimulus and the second; and Field 4 was on during the inter-cycle interval (ICI). The duration and intensity of each of the four fields were individually pre-determined. At the start of a period of observation, Field 4 was held on indefinitely. After a switch was thrown, Field 4 was extinguished, the cycle began, was repeated automatically as many times as desired, and was stopped manually with Field 4 again illuminated. The duration of Fields 1, 2, or 3 could be changed during the on-time of Field 4, even when the device was recycling uninterruptedly.

The optical apparatus was a stereoscopic tachistoscope. A top view of the box with its cover off and a side view of the right half with the outer wall removed are shown in Figs. 1 A and 1 B respectively. (For Fig. 1 A the eyes are at the bottom of the diagram looking toward FP (the fixation-point); for Fig. 1 B they are at the left looking toward FP.) The light from two tubes (L 1 in Fig. 1 B, shown for the right-hand field), which were equidistant from the middle of the stimulus card (C), fell on the card and was reflected back through a semi-reflecting mirror (M 1).

---

[15] An improved version of this device has been described by H. S. Koletsky and P. A. Kolers, A multi-field electronic tachistoscope, this JOURNAL, 72, 1959, 456-459.

[16] Similar sources have been used by others at even shorter pulse-widths. See George Humphrey, P. G. M. Dawe, and Bernard Mandell, New high-speed electronic tachistoscope, *Nature*, 176, 1955, 231-234. The latency and striking voltage of cold cathode gas tubes vary when the tubes are worked in the dark. The mechanism of this phenomenon is assumed to be related to the curtailment of the supply of photons which normally would ionize the gas when a potential appeared across the lamp. A description of several methods of reducing the variability of latency and striking voltage is being prepared. In the present case we found the maximal latency to 'full on' to be of the order of 5 m.sec. for short pulses (10 m.sec.) occurring once every 5 sec., after a considerable period of continuous operation in the dark, and to be less when the pulse width was increased. Since the variations were not observed to be systematic, we believe that no particular set of data was biassed unduly.

This light was then reflected by a front-surface mirror (M 3) through a square aperture in the head-rest (not shown) and into the right eye. In Field 3 the light reflected by the left-hand card passed through semi-reflecting mirror (M 2) and through another square aperture in the head-rest directly into the left eye. Natural pupils were used.

The light from Fields 2 and 4 (L 2 and L 4 in Fig. 1 A) was diffused by flashed opal glass (G) and reflected by mirrors (M 1 and M 2) through the apertures in the viewing hood and into the eyes. Slight differences in color between the light from the stimulus-fields (Fields 1 and 3) and the fields filling the pauses between stimuli (Fields 2 and 4) were found in this arrangement. These differences were

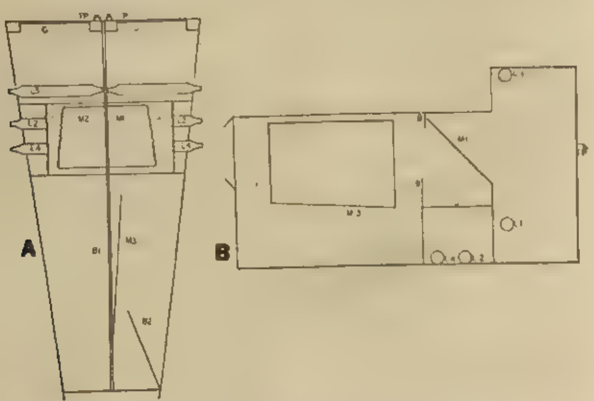


Fig. 1. Viewing Apparatus

(A) Top view of stereoscope, showing only one of two light tubes in Fields 1 and 3. The fixation-light to the right eye was just visible between the edges of (M 3) and (B 2).

(B) Side view of right half; baffle B2 not shown. FP = fixation-points; C = stimulus-cards; L 1, L 2, etc. = light tubes in respective viewing fields; G = flashed opal glass; M 1, M 2 = semi-reflecting mirrors; M 3 = front-surface mirror; B, B 1, B 2 = baffles.

caused by the semi-reflecting mirrors which transmit and reflect wave-lengths differentially. Fry's experiments, however, suggest that these slight color-differences are probably irrelevant in the present context, and we have proceeded on this assumption.

Two red fixation-points (FP), approximately 1 mm. in diameter, were located at the rear of the viewing box at a distance of 67.5 cm. from the corneas. These were fixated and fused foveally. A baffle (B 2) placed near the viewing hood for the right field prevented a direct view of the stimuli in that field by the right eye; the right field was seen only in the mirror. The left field was seen directly by the left eye. When the fixation-spots were fused, the virtual image presented by mirror (M 3) and the object seen by the left eye appeared concentric and to the left of the fixation-point. The entire device was lined throughout with black flocked paper.

Field 1 stimulated only the right eye and Field 3 only the left eye; on the other hand, Field 2 stimulated both eyes. Accordingly, each half of Field 2 was made

approximately half the intensity of Field 1 (or 3). Ten measurements of Field 1 and Field 3, made through the apertures in the viewing hood with a Macbeth Illuminometer, gave a mean reading of 1.74 ft.-c. Each half of Field 2 was made approximately half this value, and Field 4 was treated in the same manner. The head-rest viewing hood was itself padded with sponge rubber and contoured to fit snugly around the upper half of *S*'s face. Ambient light was thus blocked from reaching the eyes.

Stimulus-forms. The forms used in these experiments were usually disks and rings made with drawing ink on heavy cardboard. The forms were either black on white

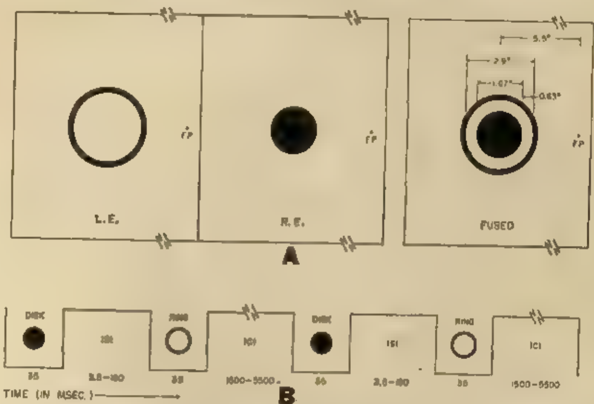


FIG. 2. (A) APPEARANCE OF THE DISK AND A LARGE RING ARRANGED FOR THE DISK-RING SEQUENCE

L.E. is stimulus-form for left eye; R.E. for right eye; "Fused" is appearance of the figures when lighted simultaneously. The areas above and below the fixation-points (FP) were black.

(B) DIAGRAM SHOWING TEMPORAL RELATIONS DURING TWO CYCLES OF PRESENTATION

The initial light-adapting condition is not shown.

cardboard or white on black cardboard. The first figure in a sequence was always presented to the right eye and the second to the left. The visible fields appeared coincident and measured 8.2 × 9.7 cm. each.

The figures appeared in the middle of the lighted part of the visible fields. The mid-point of the figure presented to the left eye was 5.5° to the left of the fixation-point. The forms were made to appear concentric by manipulating their position to the right eye or of the front-surface mirror (M 3) by which it was reflected.

In all cases the thickness of the ring-wall was constant at 4 mm. (0.33°). Most experiments were performed with a disk of 20 mm. diameter (1.67°) and rings whose inner diameters ranged from 20 to 40 mm. (1.67–3.35°). A few experiments were performed with disks of small diameters and correspondingly smaller rings. Unless otherwise indicated, the results presented are for the 20-mm. disk and

appropriate rings. All calculations of visual angle were based upon a distance from the stimulus to the cornea of the left eye (67.5 cm.). Fig. 2 A illustrates the appearance of the forms, and Fig. 2 B gives the temporal course of a presentation.

Training. We found, after five or six practice-sessions, that the range of variation in means of dichoptic data was usually greater than 30%. Protracted training was undertaken with Kolers as *S* in an effort to reduce the magnitude of this variation. The data presented here were obtained after several months of practice, when variation of means had been reduced to about 15% between sessions and 10% within sessions. While high, these numbers are not atypical for dichoptic observation.[17] From time to time other *S*s were trained for varied periods. In all cases the shapes of the functions obtained with these latter *S*s were the same as those obtained from Kolers, although absolute values differed.

Instructions. To simplify the response-situation *S* was instructed to tap the table twice if he saw the figure whose discriminability was being investigated, and to tap it once if he saw only the other figure. He was carefully instructed that the discrimination was between certainty and uncertainty; if he was unsure whether he had seen the relevant figure, his response was to be a single tap; only if he was sure he had seen the figure was he to tap twice. After practice the *S*s reported no difficulty with this method of responding and it was used throughout the investigation.

Procedure. The variability of Alpern's data, obtained with the method of limits, appears to go as high as 44%. We therefore investigated several other psychophysical methods in a search for one which would yield more reliable data for the difficult discrimination the *S*s were required to make. This search indicated that if a modification of the quantal method were used in which only five responses were made at a given stimulus-setting, sequential dependencies among responses remained reasonably low and the data were fairly reliable.[18] Even with this method, however, the data exhibit marked 'anchoring effects' as a result of the range of stimulus-conditions used. Fig. 3 shows an example in which three different ranges for the duration of Field 2 were employed. As a result, the probability of seeing the disk, p (D), also changed. Thus, the amount of masking obtained at any combination of stimulus-conditions was found to vary with the range of conditions explored.

Throughout this investigation, five reports to a given value of the independent variable were followed immediately by five reports at each of four other values of the independent variable. A 'set' of data then consisted of 25 reports, 5 for each of 5 values of the independent variable. To reduce sequential dependencies

[17] See, for example, C. G. Mueller and V. V. Lloyd, Stereoscopic acuity for various levels of illumination, *Proc. Nat. Acad. Sci. Wash.* 34, 1948, 223-227; J. F. Schouten and R. L. Ornstein, Measurements on direct and indirect adaptation by means of an indirect method, *J. opt. Soc. Amer.*, 29, 1939, 168-182.

[18] The quantal method is described by S. S. Stevens, C. T. Morgan, and John Volkmann, Theory of the neural quantum in the discrimination of loudness and pitch, this JOURNAL, 54, 1941, 315-335. The occurrence of sequential dependencies between responses with this method is discussed by Neisser, among others. Ulric Neisser, Response-sequences and the hypothesis of the neural quantum, this JOURNAL, 70, 1957, 512-527.

among parts of the set, two presentations at a given stimulus-setting were made before the first report was required, hence seven presentations were made at each value of the independent variable. Several sets of data were collected in a session; characteristically, a set for each of five values of the parameter. For practice the high, low, and mid-point of the range of the independent variable were presented a fixed number of times after a change in the parameter.

An experimental session had the following form: (1) an initial 2 min. of dark adaptation; (2) 30 sec. of light adaptation to Field 4; (3) 30 sec. of practice; (4) about 100 sec. collecting data; (5) 1 min. of rest with eyes closed; (6) a repetition of the sequence starting again with (2) but with another value of the

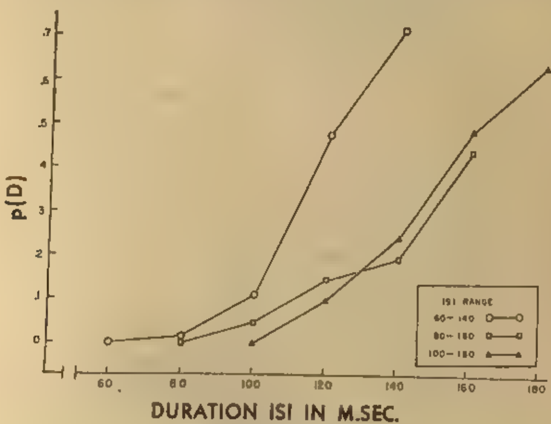


Fig. 3. Anchoring Effect of Range of ISI

Probability of seeing the disk, p(D), in disk-ring sequence, as a function of duration of the inter-stimulus interval (ISI). Three experiments, each with stimuli in the range shown. Disk and ring presented for 35 m.sec. each; inter-cycle interval (ICI) was 1500 m.sec.

parameter. Not uncommonly during a session the fixation-points blurred and the forms did not appear concentric. When this occurred, *S* did not respond to the presentations until the points again appeared sharp and the forms concentric.

In the results reported below, data from occasionally two but usually four sessions are combined in a single curve. Thus each point plotted represents either 10 or 20 reports gathered from counter-balanced random orders of presentation. After considerable practice by *S* we found that two sessions gave fairly reliable data if a rest-period of at least a half-hour in a lighted room intervened between sessions and if fewer than 600 observations were made on a given day. Choice of the order in which values of each variable were presented, the manipulation of apparatus and recording of the data were performed by various assistants.[10]

---

[10] Of the several people who acted as *Ss* or as assistants, we wish particularly to thank Mrs. Bettie Lou Nims and Messrs. Ronald Bray, Richard Peckham, and Alexander Stern.

Results

(1) Masking as a function of ICI. Werner found the masking effect to be maximal at inter-cycle intervals (ICIs) between 280 and 560 m.sec. Since even the latter is a relatively short time, it is possible that some part of the inhibition of the disk is due to an action of each presentation of the ring upon *subsequent* presentations of the disk. Fig. 4 A shows that the probability of detecting the disk, p(D), changes as a function of the ICI. Four curves are shown, for ICIs from 1500 to 5500 m.sec. Each curve represents the proportional frequency with which the disk, the first of the two figures, was reported. In the range employed here of the inter-stimulus interval (ISI), the pause between the disk and ring, p(D) increases

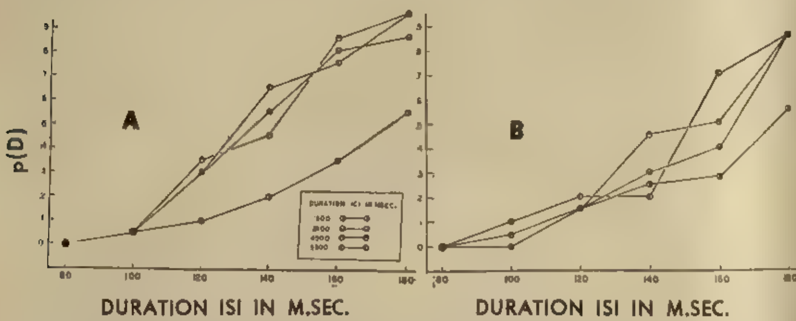


Fig. 4. Effect of 'Fatigue' in Successive Sessions on One Day
(A) First two sessions. Probability of seeing disk, p(D), as a function of ISI and of ICI. Disk and ring presented for 35 m.sec. each, separated by ISI given on abscissa. Disk of 20 mm. diameter; ring of 20 mm. inside diameter. Twenty observations per point. (B) Second pair of sessions.

regularly with an increase in ISI. Interest at the moment, however, is in the differences among the four curves. It can be seen that, at an ICI of 1500 m.sec., p(D) is less than at ICIs between 3500 and 5500 m.sec. Data were also taken, but no curve is shown, for the ICI of 2500 m.sec. The variability of the data at this ICI was three times greater than that for the other ICIs. Subjectively it seemed that a 2500 m.sec. pause was too long for sustained vigilance and too short to permit relaxation betwen presentations.

Each point on the curves of Fig. 4 A is based upon a total of 20 observations which were made over 2 days in groups of 5, as described earlier. Fig. 4 B is based upon observations, the same number per point, made after a 15-min. rest following collection of the data for Fig. 4 A. As can be seen in Fig. 4 B, increasing the number of observations made on a given

day increases the amount of masking (inhibition of the disk) at the longer ICIs, making the curves for the latter approach that for the shortest ICI. The two curves for the 1500 m.sec. ICI (Figs. 4 A and 4 B) are almost identical.

These graphs might suggest that masking at the longer ICIs increases with the number of observations made and approaches a maximum at the 1500 m.sec. ICI. Actually, however, a further shortening of the ICI to less than 1500 m.sec. results in additional inhibition of the disk. Under conditions that yield a $p(\mathrm{D})$ of 0.90 with a long ICI, probabilities of detection of 0.30 or less can be obtained by shortening the ICI to the values Werner used. Clearly, there is an inhibitory action of a ring upon the *next* presentation of the disk when the temporal separation between cycles is short enough. To minimize this proactive component of masking, all data were subsequently collected with the ICI at 3500 m.sec.

(2) Masking as a function of ISI. The results presented in Fig. 4 show that the probability of detecting the disk, $p(\mathrm{D})$, falls off as the interstimulus interval (ISI) is decreased from 180 to 80 m.sec. As the ISI was still further shortened from 80 to 40 m.sec., $p(\mathrm{D})$ remained zero. At an ISI of approximately 30 m.sec., however, a change occurred: while the disk itself was still not discriminable, the ring appeared to be much thicker and heavier. With further shortening of the ISI to less than 20 m.sec., the disk became visible occasionally, and disk and ring then appeared together as a filled ring (or larger disk). As the ISI was shortened still further, the disk continued to be seen, sometimes as a whole figure, but most frequently as an irregular one, and occasionally with greater contrast, than at the longer ISIs.[20]

These observations suggest that the full masking effect requires a fixed time for its development: holding durations of the disk and ring constant, an orderly sequence of appearance, disappearance, reappearance occurs as the ISI is increased from its shortest to its longest value.[21] This sequence is shown in detail in Fig. 5. Efforts were made to record this sequence in a single experimental session. To do this required considerable lengthening

[20] When the inner diameter of the ring was greater than the diameter of the disk, the disk almost never appeared deformed at these ISIs. The greater contrast mentioned, in which the disk appeared blacker when followed by the ring than when presented alone, occurred intermittently. With our apparatus we were not able to establish the requirements for the occurrence of this apparent enhancement.

[21] The analog to this in Alpern's data is a zero brightness difference between his test-flash and a standard at the shortest ISIs. This is shown for the less intense stimuli in his Fig. 4. Alpern, *op. cit.* 1953, 650.

of the session, and the data became more variable as the number of observations increased. Consequently, separate sessions of observation were carried out, and the results for two different ISI ranges are presented, from 3.5 to 80 m.sec. in Fig. 5 A, and from 60 to 180 m.sec. in Fig. 5 B. The anchoring effects shown earlier (Fig. 3) and, possibly, the fact that the data were collected some weeks apart, may explain the lack of better coincidence between the two sets of curves in the region in which they overlap.

It can be seen in Fig. 5 A that at the shortest ISI the disk is usually dis-

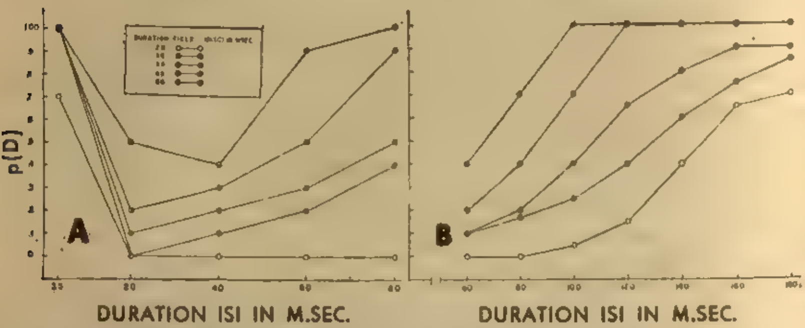


FIG. 5. EFFECT OF DURATION OF DISK AND OF ISI

Probability of seeing disk is shown as a function of ISI with duration of disk as parameter. (A) Shorter durations of ISI. Ring just circumscribes 20 mm. disk. (B) Longer durations of ISI. Duration of ring was 35 m.sec.; ICI was 3500 m.sec.

criminable, although frequently it appears deformed. Even with foveal observation at the shortest ISIs, a disk apparently exactly contiguous to the inner wall of the ring will usually appear to have a faint gray border or to be smaller than the ring. For purposes of the experiment these various distortions in the appearance of the disk were disregarded and *Ss* were instructed to report that they had seen the disk if a major part of it was discriminable.

The distortions, and thus the frequency of report at the shortest ISIs, may, however, have been the result of artifacts. The first possible source of error concerns the light-sources. Although the latencies to 'full on' and 'full off' were 1.4 m.sec. each, the shape of the transition is not the same in the two cases, for the onset of the tube is sharper than the decay. It was conceivable then that light from a decaying field added to light from the succeeding field coming on, and that this summation lowered the threshold for seeing the disk. This possibility was investigated in a control experiment. The threshold for the disk in Field 1 was determined when it was followed by a plain white card, rather than a ring, in Field 3 at ISIs between 3.5 and 80 m.sec. Under these conditions p(D) was found to be about 10% higher at

the ISI of 3.5 m.sec. than at ISIs longer than 20 m.sec., and with no difference between 20 and 80 m.sec. This 10% difference in p(D) is obviously not large enough to account by itself for the high p(D) at the 3.5 m.sec. ISI of Fig. 5A.

A second possibility which occurred to us as a source of the high p(D) at the short ISIs was that the light of Field 2 at the short durations of that field in some way served to enhance p(D). Control studies, however, showed that the disk followed by a contiguous ring is still visible at the 3.5 m.sec. ISI even when the lamps of Field 2 are turned completely off, so that the field is of zero intensity. In this condition, as with a lighted Field 2, the sequence of appearance, disappearance, reappearance occurs.

Therefore, if the above artifact were eliminated at the shortest ISIs, p(D) might

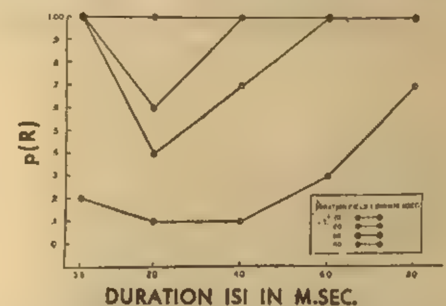


FIG. 6. EFFECT OF DURATION OF RING AND ISI ON RING-DISK SEQUENCE
Disk was presented for 35 m.sec.; ICI was 3500 m.sec.

be slightly less than is shown in Fig. 5 A. But it appears certain that a U-shaped curve correctly describes the course of masking as a function of ISI.

(3) Reversal of sequence: Ring followed by disk. In the explanation of his data Werner implied that a ring followed by a disk was *not* inhibited at the durations he used because its development required a shorter time than did the disk. The data of Fig. 6, compared with those of Fig. 5 A, support this suggestion. The shapes of the curves describing the inhibition of the first figure are maintained in the ring–disk sequence, but the ring, in this condition, must be of very brief duration if it is to be masked. Stated in other words, a two-bordered form inhibits the appearance of an enclosed one-bordered form more than the latter does the former.

A comparison of Fig. 5 A with Fig. 6, and a further comparison with Fig. 7 below, reveals some uncertainty about the locus of the minimal point of the U-shaped curves. It is possible that the ISI at which maximal masking occurs varies slightly as inner diameter of the ring or duration of the first figure is changed. This variation, if it occurred, would not be seen because of the large ISI steps used.

(4) Masking as a function of angular separation: As noted earlier, Werner had found that absolute contiguity of the borders of the disk and ring was a necessary condition for masking to occur. We found, however, as Alpern and Fry did with flashes of light, that masking will still occur with some separation of the borders of the forms. The amount of masking that a second form can exert on the first is inversely related to the amount of visual angle that separates the two forms. This is shown in Fig. 7 for masking situations in which rings of varied diameters followed a disk of 20 mm. diameter. Some masking occurs even with the disk 0.63° from the inner wall of the ring; this is the difference in subtense of the radii of a disk of 20 mm. diameter and a ring of 34 mm. inner diameter. Even the

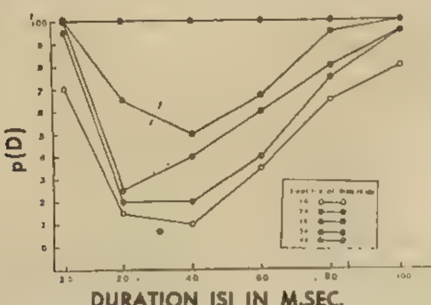


FIG. 7. INHIBITION OF DISK BY RING AT VARIOUS ANGULAR SEPARATIONS
Inhibition consists of decreased probability of report, p(D). Disk and ring presented for 35 m.sec. each; ICI was 3500 m.sec.

largest ring (40 mm. inner diameter) can inhibit the disk if the ICI is shortened sufficiently.

The data presented in Fig. 8 show the same relationship but with a smaller disk. The shapes of the functions in the two cases are comparable.

One other thing is of interest in Fig. 8. This is that maximal inhibition occurs at longer ISIs as the separation between the forms is increased. This finding is not in agreement with Alpern's. In his Figs. 10 and 11 he shows that maximal inhibition occurs at *earlier* ISIs with more widely separated figures.[22]

(5) Foveal effects. To examine whether masking can occur in the fovea, disks of three different diameters were used: 20, 10, and 5 mm. (1.67, 0.84, and 0.42° respectively). In place of the fixation-points, two red dots

[22] Alpern, *op. cit.*, 1953, 653 f.

about 0.5 mm. in diameter were painted onto the flashed opal glass in Fields 2 and 4, one on each side of the mid-line baffle, which *S* learned to fuse foveally. The disk and ring were so arranged that their mid-points appeared to coincide with the red dots.

Under these conditions we found that masking tended to diminish as the stimulus-objects were made physically smaller. A ring of inner diameter equal to the diameter of the disk inhibited the appearance of the disk at some given ISI. But the smallest angular separation between the border of the 5-mm. disk and a ring served to make the disk again visible. Thus, while foveal masking occurs, it is very much less extensive than masking in the periphery.[23] Even, however, with a disk of 5-mm. diameter, the visual

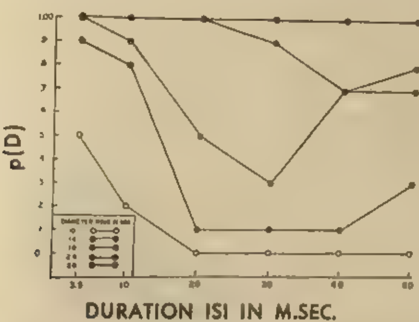


Fig. 8. Effect of Angular Separation with Smaller Disk and at Shorter ISI-Steps

Diameter of the disk was 10 mm. (0.84°). Disk and ring presented for 35 m.sec. each; ICI was 3500 m.sec.

angle occupied (0.42°) is greater than the extent of the *fovea centralis*,[24] hence it still is possible that no masking effect will be found in that region. We were unable to test this possibility with our apparatus.

(6) Paracontrast. The results presented in Fig. 4 A imply that the appearance of a figure can be inhibited by flashing another figure before it. This type of inhibitory action is sometimes called paracontrast. The occurrence of this action was apparently unknown to Werner. Alpern, however, believes that it is due to the residual action initiated by intra-ocular stray light following the exposure of the first flash. If this belief were cor-

---

[23] In Alpern's experiment there was at least 0.25° separation between the borders of his test- and inhibiting-figures which, in the light of the above results, could account for his inability to record any metacontrast effects in the fovea (*Ibid.*, 649).

[24] S. L. Polyak, *The Vertebrate Visual System*, 1957, 1-1390.

rect, paracontrast would be a phenomenon different in kind from metacontrast, and so Alpern has maintained.

We have found, however, that the functions for paracontrast tend to complement those of metacontrast: as the duration of the first figure is increased, the probability of seeing the second decreases; as the angular separation between figures is increased, the probability of seeing the second also increases. Further, we have found that in the ring–disk sequence the ring exerts more inhibition upon the disk than the disk, in the reverse sequence, exerts upon the ring. Fig. 9 shows the results of an experiment in which this forward-acting inhibition was measured. The disk-ring sequence was used and the inhibitory action that the disk exerts on a suc-

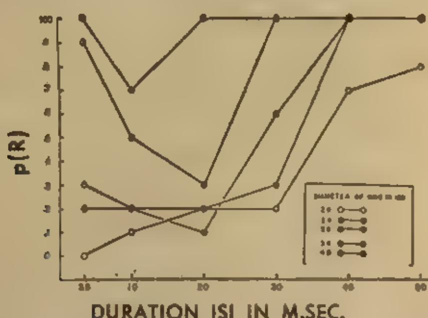


Fig. 9. Paracontrast: Proactive Effect of Disk on Ring in Disk—Ring Sequence as a Function of ISI and Size of Ring

Disk of 20-mm. diameter was presented for 35 m.sec.; rings presented for 7.5 m.sec. each; ICI was 3500 m.sec.

ceeding ring as a function of angular separation of the two figures and of the ISI is shown. In collecting these data the disk was presented for 35 m.sec. and the ring for only 7.5 m.sec. Since in our experimental arrangement only one form was flashed to each eye, the paracontrast effect cannot be attributed to residual effects of stray light in the eye. Rather, it seems to be an inhibitory effect intimately related to masking and metacontrast. Each of the two forms in a sequence affects the other at the durations studied, but the inhibitory action of a form when it is second is greater than when it is first.

(7) Dichoptic effects. Since we encountered no great difficulty in obtaining dichoptic masking with disk and ring as stimulus-figures (Figs. 3–8), it became desirable to explore the reasons for Alpern's failure to find an equivalent effect. To this end a different set of stimulus-figures was constructed, and with this set no masking was observed at any duration

of the figures or of the interval between them. This set consisted of a black rectangle (b), 25-mm. high × 6-mm. wide as the first figure presented to the right eye; and two rectangles (c–c′), each of the same dimensions, as the second figure presented to the left eye. In shape and angular subtense these were somewhat similar to the figures Alpern used, although his presentations were of light against a dark ground. The single rectangle (b) was so positioned that it appeared to be flanked symmetrically by the two rectangles (c–c′) at various distances from the mid-point of (b). This arrangement created a total display of a single rectangle followed by two flanking ones.

If the first rectangle (b) were masked, one would see only the two flanking rectangles (c–c′) at some ISI, but the masked rectangle would gradually appear as either its duration or the ISI were increased (compare Fig. 5). What happened instead is that *S* continued to see only two rectangles over the entire range of stimulus-times explored, and usually one of them appeared much darker than the other. It appeared that a summative, rather than an inhibitory, effect occurred between the single rectangle (b) of the first presentation and one of the pair of rectangles (c–c′) in the second presentation. Alternatively, depending upon the duration of the ISI, apparent movement was perceived, either in one direction only, or from the mid-point out in both directions.

Such observations made it seem likely that the failure to observe masking with these figures was due to eye movements or to improper fixation that tended to alter the relative positions of the figures. Although the first figure was not seen, the cause of its absence would appear to be quite different from that in a masking situation.[25] We then reasoned that, if eye-movements were the sufficient condition for the failure to observe dichoptic masking or metacontrast with rectangles as stimulus-figures, two opposing processes were occurring in Alpern's experiments: a summative one between the first and second presentations when the fixation at the moment of exposure made (b) and one of the rectangles of (c–c′) overlap, and an inhibitory one when (b) appeared between the two rectangles (c–c′). In our observations the latter inhibitory effect appeared to be of greater magnitude than the former summative one. Since Alpern's measure was the difference in apparent brightness between the (b) rectangle and a standard, the net difference in brightness between (b) and the standard taken over a series of presentations would have been close to zero, but the *variability* of brightness matches would have been very large. We believe,

---

[25] The appearance that the disk and ring were concentric varied slightly, requiring occasional manipulation of the stimulus-forms to maintain coincidence.

therefore, that Alpern did not record any dichoptic metacontrast because on some presentations the apparent brightness of (b) was less and on some presentations it was greater than that of the standard.

To check this belief we prepared white rectangles on black cards and reduced the intensity of Fields 2 and 4 (the ISI and ICI) to zero, which approximated Alpern's conditions. However, instead of making the single rectangle of the first card, (b), appear on a level with the two rectangles of the second card (c–c′), we displaced (b) upwards so that half of it extended beyond the tops of the two rectangles of (c–c′). With these stimuli it was very easy to observe changes in apparent locus of (b). Approximately one-third of the time (b) appeared in the middle of the two rectangles (c–c′) and two-thirds of the time it appeared displaced to either side. When (b) appeared to be between (c–c′), the bottom half of (b) was inhibited, appearing much less bright than the top half: dichoptic metacontrast was observed to occur with flashes of light when the apparent loci of the stimulus-figures were appropriate to its occurrence.

Discussion

The present study was undertaken to resolve the discrepancy in previous reports regarding the existence of dichoptic masking and to quantify some of the parameters associated with this effect. Three of the parameters investigated, retinal locus of stimulation, shape of the figures, and paracontrast effects, yielded primarily qualitative results. Of particular interest is the demonstration that the action of two successive figures can be reciprocal: the first of two figures will inhibit the second at the same time that the second inhibits the first. Comparison of Fig. 6 with Fig. 9 shows, however, that 'retroactive' masking effects are far more potent that 'proactive' paracontrast effects. The hypothesis that paracontrast effects are the result of stray light following a flash to a dark-adapted eye is also contradicted, since paracontrast was found with dichoptic stimulation to light-adapted eyes. Both conditions are incompatible with the stray-light hypothesis.

The four other parameters described, the interval between cycles (ICI), the duration of the first figure, the angular separation between figures, and the interval within a cycle between the two figures (ISI), were investigated in enough detail to permit some quantitative statement of their relations. These statements will be restricted here to the disk–ring sequence. The probability of seeing the disk p(D) is greater after a longer ICI, *ceteris paribus;* and it appears likely that the ring presented on trial n plays the pre-

dominant role in this proactive inhibition of the disk on trial $n + 1$. (Whether the proactive effect observed when ICI is varied is attributable to the same mechanism governing the proactive effect of paracontrast is not known at present.) It should be noted that the ICI we used most often (3500 m.sec.) is more than six times longer than the longest ICI Werner used. This difference in ICI is by itself almost sufficient to explain the large difference between the ISI at which Werner and we found p(D) to be a minimum.

The probability of seeing the disk, p(D), is greater when the disk is exposed for a longer time and is separated spatially by a greater distance from the inner wall of the ring. The temporal locus of maximal inhibition (the ISI at which p(D) is minimal) varies, however, with the duration of the disk and with the angular separation between the forms. Whatever the inhibitory mechanism is that mediates these effects, it apparently takes longer to reach its maximum when the space between the figures or the duration of the first figure is increased.

The effect of ISI on p(D) is complex in that p(D) is a U-shaped function of that variable. This fact is not easily reconciled with Werner's hypothesis that masking is the appropriation of a contour. That hypothesis would seem to require that p(D) should be least at the shortest ISIs, when the disk has had least time to form.

As an alternative to the hypothesis of contour-appropriation we attempted to account for the phenomena of masking as a disinhibitory process, taking as a model the recent work by Hartline and his co-workers.[26] With this model we accounted for many of the data for metacontrast, the inhibitory effect resulting from flashes of light to the dark-adapted eye.[27] It is difficult, however, to account for the U-shaped curve of masking (the inhibitory effect resulting from flashes of black forms to the light-adapted eye) since a model based only on disinhibitory processes would predict that masking is a monotonically decreasing function of ISI. The appearance of the disk at the shortest ISIs may be the result of a facilitory mechanism, or it may be due to the relatively poor temporal resolving power of the

---

[26] H. K. Hartline and P. R. McDonald, Light and dark adaptation of single photoreceptor elements in the eye of *Limulus*, *J. cell. comp. Physiol.*, 30, 1947, 225-253; H. K. Hartline, H. G. Wagner, and Floyd Ratliff, Inhibition in the eye of *Limulus*, *J. gen. Physiol.*, 39, 1956, 651-673; Hartline and Ratliff, Inhibitory interaction of receptor units in the eye of *Limulus*, *J. gen. Physiol.*, 40, 1957, 357-376.

[27] Suggestions of a similar sort have been made in a paper which has appeared since this report was prepared. See W. S. Battersby and I. H. Wagman, Neural limitations of visual excitability. I. The time course of monocular light adaptation, *J. opt. Soc. Amer.*, 49, 1959, 752-759.

eye.[28] We are currently exploring this question experimentally. In either case, the suggestion is present that while *responses* to flashes of light and to flashes of black mirror each other, the underlying mechanisms do not.[29]

Finally, are these inhibitory effects to be attributed primarily to nervous or to photochemical action? The latter would assume a delay in nervous action associated with the period required for photochemical changes in the eye; but, as Alpern has already noted and as shown in Fig. 5 B, masking and metacontrast effects continue to occur beyond the period of a critical duration.[30] Additional evidence that the effects are essentially nervous in character comes from two sources. First, in exploratory studies we stimulated the eye electrically to evoke phosphenes. This method probably bypasses the retinal receptors entirely.[31] Nevertheless, metacontrast effects were observed when *S* was required to judge which of two successive, equal pulses produced a phosphene that appeared brighter. Metacontrast was found when the two pulses were delivered sequentially to the same eye, and when one was delivered to one eye and the second to the other eye, although the time course was different in the two cases. The second source of evidence is studies of other modalities. Miller has demonstrated 'metacontrast' in audition; there, the magnitude of the effect seemed to depend upon the relative frequencies of the two tones, although the effect seemed to disappear totally at ISIs greater than 8 m.sec.[32] The inhibition of the first of two brief, asynchronous stimuli may be a general characteristic of sensory functioning, although it occurs over far longer ISIs with photic stimulation than with other kinds.

Summary

The appearance of a black disk is inhibited when it is followed by a concentric black ring. The extent of this inhibition, which is called visual masking, was studied systematically as a function of time between cycles of observation (ICI), duration of presentation of the disk, duration of the

---

[28] G. B. Arden and R. A. Weale, Variations of the latent period of vision, *Proc. Roy. Soc. Lond.*, 142B, 1954, 258-266.

[29] See, for example, H. B. Barlow, R. Fitzhugh, and W. S. Kuffler, Change of organization in the receptive fields of the cat's retina during dark adaptation, *J. Physiol.*, 137, 1957, 338-354.

[30] Some of Alpern's data and our own preliminary observations suggest that the first form will be inhibited only if its duration is less than the period of a critical duration. See especially Figs. 6 and 7 in Alpern, *op. cit.* 1953, 651 f., and Fig. 5 of the present paper.

[31] Harry Grundfest, Electrical inexcitability of synapses and some consequences in the central nervous system, *Physiol. Rev.*, 37, 1957, 337-361.

[32] R. L. Miller, Masking effects of periodically pulsed tones as a function of time and frequency, *J. acoust. Soc. Amer.*, 19, 1947, 798-807.

pause between the disk and ring (ISI), and angular separation of the two figures. A related phenomenon is paracontrast, the inhibition of the second figure by the first. This was also studied as a function of the angular separation of the figures and of ISI.

We found that inhibition of the first figure increases with a decrease of its duration, with shorter ICI, and with less angular separation between the figures. The effect of ISI was found to be U-shaped in both the disk–ring and ring–disk sequences; that is, greatest masking occurred over a limited range of ISI, and decreased outside of this range. Data obtained on the paracontrast effect were found to complement those on the masking effect, indicating that the former cannot be due to the action of intra-ocular stray light. In comparing masking with paracontrast, it was shown that the inhibitory action of a figure is far greater when it is presented second than when it is presented first.

It was also shown that the inhibitory effect occurs in the fovea, but only when the two forms are contiguous; and that disjunctive eye movements are the probable cause of the difficulty in finding inhibitory effects in dichoptic observation.

Finally, evidence from other modalities and from direct electrical stimulation of the eye is offered to suggest that the kind of inhibitory effects studied here may be characteristic of all sensory functioning.

REPETITION AND ASSOCIATIVE LEARNING

By Lester L. Clark, Theron G. Lansford, and Karl M. Dallenbach, University of Texas

In a recent number of this Journal, Rock reported the results of a study on the role of repetition in associative learning, which, if corroborated, is of great theoretical importance.[1] Using the method of paired-associates, he found, in two series of experiments—the first with 12 letter-number pairs and the second with 8 three-letter pairs of nonsense-syllables—that pair presented but once before learning were learned as rapidly as those repeated until they were learned. From these results, he concluded: "that repetition plays no role in the formation (as distinct from the strengthening) of associations, other than that of providing the occasion for new ones to be formed, each on a single trial."[2]

These results stand in direct opposition to the modern as well as to the traditional theories of learning. From Aristotle through the Associationists to Ebbinghaus, Thorndike, Müller, Watson, Hunter, and Hull (to mention a few who exemplify the many), repetition or some form of it, such as frequency or exercise, has been regarded as the principal basis of learning. Ebbinghaus, for example, states that the relationship between repetition and learning "can be described figuratively by speaking of the series as being more or less deeply engraved in some mental substratum [and that] as the number of repetitions increases, the series is engraved more and more deeply and indelibly."[3]

Though some items in a series are frequently learned in a single trial, these instances are regarded by most learning theorists as special cases explained by the secondary laws of association, such as primacy, recency, vividness, effort, etc., and the series itself is regarded as being learned gradually, a bit at a time; every repetition adding an increment until the entire series is learned. These special instances did not greatly strain the role of repetition in learning, particularly as each has its own explanation, but single-trial

---

* Received for publication January 6, 1958.

[1] Irwin Rock, The role of repetition in associative learning, this Journal, 70, 1957, 186-193.

[2] *Ibid.*, 193.

[3] Hermann Ebbinghaus, *Memory*, 1885, trans. by H. A. Ruger and C. E. Bussenius, 1913, 52.

(insightful) learning, demonstrated by the Gestalt psychologist[4] and the experimenters concerned with the learning of relations,[5] would not fit into the repetitive pattern, hence the theorists either ignored it or regarded is as an insightful phenomenon outside their field of concern.

Consideration of Rock's demonstration of single-trial learning may not, however, be so easily avoided as his results fall within the realm of associative learning. If, therefore, his results are confirmed, if they are not artifacts of his method, which has still to be proved, then all of the modern-day theories of learning that are based on repetition, or some form of it, will have to be abandoned or very radically modified.

Our first task was to discover whether Rock's results could be duplicated in the atmosphere of another laboratory. When that was done, it would be time enough to determine whether his results were artifactual.

EXPERIMENT I: REPLICATION OF ROCK'S STUDY

Experiment I is a duplication of Rock's first experiment. His method and procedure, which were faithfully followed, are described here in detail.

Stimulus-cards. The stimulus-cards, 50 in number, were 6 × 11 cm. in size. They were cut from heavy, white Strathmore board that they would not bend or crush

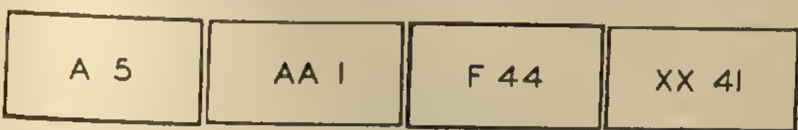


FIG. 1. SAMPLE STIMULUS-CARDS
(Cards, 6 × 11 cm.; printing, 12 mm. in height.)

when exposed in the mechanical card-changer. The letters and numbers were drawn, 12 mm. in height, upon the face of the cards with a printing set and India ink (see Fig. 1). With the exception of 'I,' which was omitted because of its similarity to the Roman numeral, all the letters, which were printed in capitals, were used twice over: once singly and once doubled. The letters were placed to the left of center; the numerals (Arabic) from 1 to 50, assigned to the letters by a table of random numbers, were placed to the right of center. Upon the backs of the cards, for use in the recall-tests, the letter or letters appearing on their faces were centered.

[4] Kurt Koffka, *Growth of the Mind*, trans. by R. M. Ogden, 1924, 214-219; Wolfgang Köhler, *Gestalt Psychology*, 1929, 371 f., 376 ff.

[5] K. M. Dallenbach, A note on the immediacy of understanding a relation, *Psychol. Forsch.*, 7, 1926, 268 f.; G. L. Kreezer and K. M. Dallenbach, Learning the relation of opposition, this JOURNAL, 41, 1929, 432-441; Mary Schooley and G. W. Hartman, The role of insight in the learning of logical relations (part-whole, action-agent, attribute-substance, genus-species, and their converse), this JOURNAL, 49, 1937, 237-292; J. I. Lacey and K. M. Dallenbach, Acquisition by children of the cause-effect relationship, this JOURNAL, 52, 1939, 103-110.

Subjects. The *S*s, 50 in number, were volunteer college students who served without knowledge of the problem. They were required to memorize 12 of the stimulus-cards by the method of paired-associates. They were divided into two groups, Control and Experimental, by chance, being assigned alternately to them as they reported to the laboratory for the experiment. No attempt was made, as in Rock's case, to equate the groups in their learning abilities.

After shuffling the cards, 12 were randomly drawn for every *S* of each group. The Control *S*s memorized the set drawn; the Experimental *S*s merely began with that set; the cards they failed during the recall-test were removed and new cards, drawn haphazardly from the pool, were substituted.

Instructions. The following instructions were given to the *S*s.

I am going to show you 12 cards in succession that have letter-number combinations upon them. You are to associate the letters with the numbers. After every exposure of the cards, you will be tested to see how many you have learned. The exposures and the tests will succeed one another until you have mastered the list. [The Experimental *S*s were, in addition, read the following.] Although 12 cards will be shown you in a series, they may not always be the same ones; new cards may be introduced in successive series. Learn all the pairs shown in the series upon which you are tested.

Procedure. The card-changer, described on p. 139, was used to expose the stimulus-cards. It was placed upon a table before *S* with the exposure-window (4.5 × 9.5 cm.) centered at his eyes. Its screen concealed *E*, the stimulus-cards, and the data-sheets and records from *S*.

Every stimulus-card was exposed for 3 sec. at intervals between successive exposures of 5 sec. The timing was controlled by an electric clock. The cards, as exposed, were stacked in order behind the changer. At the end of the learning-trial, the cards were reversed, shuffled, and presented to *S*, whose task was to call out the associated numbers. This procedure was repeated with the Control *S*s until the paired-associates had been learned. The procedure with the Experimental *S*s was different. As they responded on the recall-test, *E* removed the cards whose number-associates were not learned and substituted new cards from the pool for them. The new cards were then shuffled with those *S* had learned and the stimulus-cards were again presented to *S*. This procedure was repeated until all the letter-number associates had been learned; every pair, if learned, being learned at a single presentation.

The recall-test followed the stimulus-exposure as soon as *E* could reshuffle the cards and place them in the card-changer. An interval of 30 sec. elapsed between the recall-test and the subsequent reëxposure of the cards. During this period, *E* recorded *S*'s reports, reversed and reshuffled the stimulus-cards, and placed them in the changer. Rock's criterion of learning, which was adopted in this study, was the first correct recall of the paired-associates.

Results. Rock found, as will be recalled, that the pool of stimulus-cards was not sufficiently large for him to bring all of his Experimental *S*s to criterion. Five of them failed to reach it. To score his results, he resorted to two methods: (1) he extrapolated from the data at hand and scored the failures as 10, the number of repetitions he figured these *S*s would have required; and (2) he excluded the 'failures' from his computations. Having

placed the upped limit at 10 repetitions for his 'failures' among the Experimental *Ss*, he also placed it at that number for the three *Ss* from the Control Group who did so badly that work with them was discontinued before criterion was reached.

None of our Experimental *Ss* failed to reach criterion, but two of our Control *Ss* did. Since we were duplicating Rock's procedure, we followed his method of scoring. Table I shows the results of his and our experiment. They are as similar as experiments conducted in different laboratories can be; the differences are slight and are statistically insignificant. In both of the studies the differences between the Control and Experimental *Ss* are small

TABLE I

MEAN AND MEDIAN NUMBER OF TRIALS REQUIRED FOR CRITERION (Experiment I)

Experiment	*Ss*	Trials: All *Ss**				Trials: Successful *Ss*†			
		N	*M*	*Med.*	*SD*	*N*	*M*	*Med.*	*SD*
This one	Control	25	5.12	4.10	2.5	23	5.12	4.10	2.5
	Experimental	25	5.08	4.75	1.7	25	4.65	4.58	1.1
Rock's	Control	25	—	4.75	—	22	4.55	—	1.9
	Experimental	25	—	4.75	—	20	4.35	—	1.2

* *Ss*, who failed to reach criterion, given score of 10.
† *Ss*, who failed to reach criterion, excluded.

and statistically insignificant, but what little difference does exist is uniformly in favor of the Experimental *Ss*—*i.e.* single-trial learning.

We succeeded, in our replication of Rock's experiment, in duplicating his results.[6] Before we are justified, however, in accepting his conclusions that "repetition plays no role in the formation . . . of associations, other than that of proving the occasion for new ones to be formed, each on a single trial," we must determine whether his results (and ours) are not artifacts of the method employed. It is still a far cry from his results to his conclusions.

As Rock points out, there is one among the various conditions affecting learning in this experiment which may operate in favor of the Experimental *Ss*; namely, differences among the stimulus-cards in difficulty of learning. Some of the letter-number pairs may be learned easily, others only with difficulty; a conclusion warranted by the performance of the Control *Ss* who learned some pairs quickly and others slowly. Assuming differences in difficulty among the cards and that the 12 stimulus-cards drawn from the pool

[6] Since the acceptance of this study, two papers have been published corroborating Rock's results; see Irwin Rock and Walter Heimer, Further evidence of one-trial associative learning, this JOURNAL, 72, 1959, 1-16; and Michael Wogan and R. H. Waters, The role of repetition in learning, this JOURNAL, 72, 1959, 612-613.

will contain easy and difficult cards, then the Experimental *S*s will have an advantage. The difficult cards will be removed from their lists as rapidly as met and eventually their lists will contain only easy pairs. The Control *S*s, on the other hand, must continue repeating the pairs drawn by chance for them; the difficult together with the easy.

There was, however, no communality among our *S*s regarding the cards that were easy or difficult to learn. Neither early learning, in the case of the Control *S*s, nor single-trial learning in the case of the Experimental *S*s, clustered around certain cards. 'Success' and 'failure' were spread fairly widely throughout the cards in our stimulus-pool. This means that there are no objective differences among our stimulus-pairs in ease or difficulty of learning but it does not eliminate the subjective or idiosyncratic differences that may exist among them. Some pairs are unquestionably learned more easily by some *S*s than by others. For example: 'V 26' will be learned more easily by an *S* whose first name begins with 'V' (Victor) and is 26 yr. old than by an *S* who has neither of those mnemonic associates at his service. It may well be, therefore, that idiosyncratic factors differentiate the stimulus-pairs in difficulty and thus operate, as described above, to favor the Experimental *S*s in their task. We propose in the following experiment to investigate this possibility.

EXPERIMENT II: IDIOSYNCRATIC FACTORS

The purpose of Experiment II was twofold: first, to determine whether the idiosyncratic differences in difficulty among the stimulus-cards were responsible for the results in Experiment I—*i.e.* for the unimportance of repetition in learning; and secondly, to discover whether any other factors or conditions operated in the same direction.

It is not easy to deal experimentally with idiosyncratic factors because the only way of discovering whether a stimulus-pair is idiosyncratic is to present it to an *S;* and then it is no longer an inexperienced or unrepeated item. We avoided this difficulty, however, by the following method.

Procedure. This experiment was divided into two parts: (I) a preliminary experiment during which the idiosyncratic stimulus-items were designated individually by the *S*'s; and (II) the main experiment, conducted a week after the preliminary, during which the effect of the idiosyncratic material was noted.

Subjects. The *S*s, 28 students in psychology, were divided haphazardly into two groups of 14 each; a Control Group and an Experimental Group. None had previously participated in an experiment of this kind and none knew the purpose of this study.

Part I: Procedure. In Part I, the same materials (12 letter-number stimulus-cards), procedure, and instructions were used with the Control and Experimental *S*s as in

Experiment I. Instead, however, of drawing the stimulus-cards haphazardly from the pool, as was done in Experiment I, the cards were selected by planned haphazard and so arranged that every one in the stimulus-pool was used approximately the same number of times. *E* recorded the order in which the stimulus-cards were presented to *S* at each learning-trial, the order in which the cards were presented in the recall-tests, *S*'s responses to every card, and his remarks and observations during and after the experiment.

Learning- and recall-trials were repeated, at the temporal intervals used in Experiment I, until the criterion of one perfect recall was reached by the *S*s of both groups. The first 12 stimulus-cards that every Experimental *S* failed to learn at a single trial were removed, recorded, and used in Part II of this experiment as his stimulus-cards. These cards were assumed, since they were not learned at a single trial, to be more difficult for idiosyncratic reasons than those that were learned. They had all been seen once, an attempt had been made to recall them once, and all had been followed at least once by cards, substituted for them, that masked their memory-traces.

To obtain similar stimulus-material for the Control Group, every *S* was given, without comment or explanation, immediately after he had reached our criterion of learning, a single presentation and recall-test of 12 new stimulus-cards. After the recall-test, a second list of 12 new cards was presented. The presentations of these supplementary lists and the recall-test of the first were made at the temporal intervals used in the main experiment. The first of these two supplementary lists was the material to be learned by the Control *S*s in Part II. Like the stimulus-cards selected for the Experimental *S*s, all these cards had been seen once, had been recalled, or an attempt had been made to recall them, once, and had all been followed once by cards that masked their memory-traces.

Results. Besides supplying the stimulus-materials for Part II of this experiment, Part I further corroborated Rock's results and ours in Experiment I. Every *S* in both groups reached the criterion of learning. None, as in Rock's and our first study, failed. The mean and median numbers of trials required by the Control Group to reach criterion were 4.9 and 4.5, respectively, with an *SD* of 2.0, and by the Experimental Group, 3.9 and 3.8, respectively, with an *SD* of 1.1 (see Table II). The difference in performance between the two groups is not statistically significant but, as the table shows, all of the statistics lie again in favor of the Experimental *S*s, *i.e.* in favor of single-trial learning.

The two groups of *S*s made approximately the same number of 'failures' during the first recall-test of the main experiment—Control, 117; Experimental, 118—but after that, in all of the subsequent recall-tests, the Control *S*s made more failures (130) than the Experimental *S*s (82). The difference lies again in favor of single-trial learning.

As the opportunity was at hand, we again examined the data to discover whether there was any communality among the *S*s regarding easy and difficult stimulus-cards. We found none. The Control *S*s had little opportunity

to demonstrate this, as usually only three (never more than four) of them used stimulus-cards in common. The Experimental *S*s had, however, greater opportunity as the duplication of stimulus-cards among them was much greater. Their failures in the recall-tests necessitated the introduction of new cards from the stimulus-pool, whereas the failures of the Control *S*s did not. Neither group, however, gave any indication that the successes or failures piled up at any given card or group of cards. Both success and failure, in confirmation of the results in Experiment I, were widely and fairly evenly distributed among our stimulus-cards. If some cards were more difficult or easier than others to learn, the reasons must be idiosyncratic.

Some of the *S*s in both groups followed 'successes' with 'failures,' *i.e.* after

TABLE II
MEAN AND MEDIAN NUMBER OF TRIALS REQUIRED FOR CRITERION, ERRORS MADE DURING RECALL-TESTS, AND TOTAL NUMBER OF REVERSALS
(Experiment II)

*S*s	Part I*						Part II†		
	M	*Med.*	*SD*	errors recall-test: 1st	errors recall-test: all later	Reversals	*M*	*Med.*	*SD*
Control	4.9	4.5	2.0	117	130	32	2.9	2.8	1.2
Experimental	3.9	3.8	1.1	118	82	14	2.9	2.8	0.7

* Control *S*s repeated stimulus-cards until learned; the Experimental *S*s learned by single trial.

† Both Control and Experimental *S*s repeated stimulus-cards until learned.

learning a letter-number associate and reporting the number correctly when the letter was exposed in a recall-test, they forgot it or reported it incorrectly at a subsequent recall-test. These 'reversals' occurred 32 times for the Control *S*s and 14 times for the Experimental *S*s. Again, our results show superiority for single-trial learning.

The serial position of a stimulus-card during a presentation—which was changed haphazardly by shuffling the cards before every learning- and test-trial—had no noticeable effect upon learning. The positions of the cards learned and failed were widely and fairly evenly distributed throughout their serial positions. The 'reversals'—failures following successes mentioned above—also failed to show positional effects.

Part II: Procedure. The stimulus-cards, selected in Part I for the Control and Experimental *S*s, were presented a week later to the *S*s of the respective groups by the classical method—*i.e.* the method used with the Control *S*s alone in the previous experiments. The instructions, procedure, exposure-time and temporal intervals between the successive learning-trials and recall-tests were unchanged. The stimulus-cards of the two groups were alike, as noted above, in three respects: every *S* (1)

had seen those presented to him once; (2) had seen the letter-member of every pair in a recall-test once; and (3) had had the memory-traces of those shown him masked once by another stimulus-card. None of the cards, however, was recognized by any of the *Ss* in either group.

The stimulus-cards presented to the two groups differed, as will be recalled, in one marked respect. Those presented the Experimental *Ss* were, by theory at least, difficult—the first 12 cards that each had failed, in Part I, to learn at a single trial. Those presented the Control *Ss*, new cards selected haphazardly from the stimulus-pool, were composed, by chance, of easy and difficult cards. They should, therefore, if difficulty was the determining factor, be learned more quickly and easily by the Control *Ss* than the cards presented in the Experimental *Ss*, all of which were difficult.

Results. The results of this part of the experiment are also given in Table II. The mean and median number of trials required to bring the stimulus-series to criterion are identical (2.9 and 2.8, respectively) for the two groups. The *SD* of the mean is, however, slightly smaller for the Experimental (0.7) than for the Control *Ss* (1.2).

If idiosyncratic difficulty of the stimulus-series were responsible for Rock's and our results, then the performances of the Control *Ss* should have been markedly superior to those of the Experimental *Ss*. That they were not indicates: (1) that the failure to learn at a single trial is not a fixed or constant measure of the difficulty of a stimulus-card—that it is, at most, merely a momentary or temporary indication of difficulty which does not carry over from one occasion or situation to another; and (2) that the explanation suggested by Rock; namely, that the stimulus-lists in single-trial learning become progressively easier as the 'difficult' cards "are eliminated by a process of 'natural selection',"[7] is not correct. If it were, the Experimental *Ss* should have required many more trials to bring their 'difficult' lists to criterion than the Control *Ss* to bring their 'easier' lists to that level. That they did not indicates that idiosyncratic differences in *difficulty* among the stimulus-cards are not responsible for Rock's and our results. Search for some other factor that may be responsible for these results was, therefore, continued.

If an item of one stimulus-pair is confused during learning with an item of another stimulus-pair, the withdrawal of the stimulus-cards causing that confusion, or associative interference, from the stimulus-list facilitates learning by lessening the interference in the later learning-trials. If, on the other hand, the stimulus-cards causing the interference are left within the list, then interference may extend over several learning-trials which, by way of preservation and inhibition, will compound the difficulty of learning.

---

[7] Rock, *op cit.*, 1957, 191.

Evidence of associative interference in the results of this experiment, which we believe exists, is summarized as follows: (1) The Control and Experimental *S*s were equally free from interference when they approached the first learning-trial and they made approximately the same number of 'failures' during the first recall-test—117 and 118 respectively. (2) During the later learning-trials, as revealed by the later recall-tests, the Control *S*s, whose stimulus-cards were retained until criterion was reached, made 130 additional 'failures,' whereas the Experimental *S*s, whose stimulus-cards associated with 'failure' were removed, made but 82. The elimination of the cards causing interference reduced the number of their failures. (3) On the last recall-test before criterion was reached, the Control *S*s made approximately twice as many 'failures' as the Experimental—29 to 14, respectively. (4) The Control *S*s made more than twice the number of 'reversals' made by the Experimental *S*s—32 as against 14. When the cards causing the interference are removed, 'reversals' are decreased. (5) The variability among the Control *S*s was about twice that of the Experimental *S*s—2.0 as against 1.1, a difference of 0.9, which is significant at the 2.5% level. Variability is reduced with the elimination of interference. (6) When the two groups, in Part II, used the classical, repetitive method of learning, the mean and median number of trials required for criterion were identical, though 'easy' cards were exposed to the Control *S*s and 'difficult' cards, the first 12 each had failed to learn at a single trial in Part I, were exposed to the Experimental *S*s. (7) Interference, which spelled 'difficulty' in Part I, did not carry over to the same *S*s in Part II. The interferences, which produced 'failures,' were specific to the *S* and to the occasion, not inherent in the stimulus-cards themselves. A stimulus-card, 'difficult' for an *S* in one context, may be 'easy' for that *S* in another context. (8) The interferences were idiosyncratic as the 'failures' were evenly and widely distributed among the different stimulus-cards.

All this adds up, as we believe, to the following: (1) that, for every *S*, the learning of certain stimulus-pairs is inhibited by associative interference; (2) that the removal of those pairs from the stimulus-series facilitates learnings; and (3) that Rock's "difficulty" and our "associative interference" are not identical terms covering the same phenomena. If, however, Rock intended to include everything within his term that made learning more difficult, then the designation "associative interference" has the advantage of being specific—as we were unable to discover any other kind of idiosyncratic difficulty, and also the advantage in explaining why a specific stimulus-pair is not learned in one situation and is easily learned in another—why pairs, difficult for the Experimental *S*s in Part I, were not difficult for them in

Part II—results readily explained by the presence or absence of associative interference.

The results of Experiment II indicate that the task of the Experimental *Ss* was made easier by the removal of the stimulus-cards that caused associative interference. This factor, however, is the only one of several discussed by Rock that favored these *Ss*.[8] All the others—(1) familiarity, which comes from the repetition of the same set of cards; (2) the test, recall rather than recognition, used in measuring learning; and (3) the retention, within the learning-series, of the cards once learned and then forgotten—*i.e.* the 'reversals'—favors, as he believes, the Control *Ss*. The combined effect of these three factors should, if his deductions are correct, more than counterbalance the favorable effect of the removal of the stimulus-cards causing 'interference.' It appears, therefore, since the performance of the Experimental *Ss* were either equal to or better than those of the Control *Ss*, that repetition adds nothing to learning by single trial.

Before we draw any conclusions from the results of this experiment, we should know what the *Ss* of both groups did during the intervals in the procedure, *i.e.* (1) during the intervals between successive exposures of the stimulus-cards (5 sec.; a 3-sec. exposure plus a 2-sec. pause); (2) between the successive learning-trials and recall-tests (not timed but made as rapidly as *E* could reverse, shuffle, and replace the test-cards in the card-changer—estimated at 6–10 sec.); (3) between the successive test-cards (5 sec.); and (4) between the successive recall-tests and the next following learning-trials (30 sec.).

These intervals, which we used in duplicating Rock's procedure, are not usually long in the method of paired associates.[9] They are, however, excessively long in a study devoted to discovering the effect of repetition on learning, as they may be filled by rehearsals and repetitions. If so used, *one* learning-trial may not be counted as *one* repetition because it is, in fact, an indeterminate number. What may we conclude about repetition in such cases? The long exposure-intervals may be filled by attempts to form meaningful associations which, if made, have little or nothing to do with repetition except as repetition gives them a chance to be formed. How the *Ss* fill these intervals is very important, therefore, in the interpretation of our results.

Ss' comments. When the *Ss* had completed the experimental work, they were asked to report how they approached their task and how they had

[8] Rock, *op cit.*, 1957, 190 f.

[9] M. W. Calkins (Association, *Psychol. Monogr.*, 7, 1896, No. 2, 46 f.), who devised the method of paired-associates, used exposures of 4 to 6 sec. and pauses between successive exposures of 4 to 6 sec.

accomplished it. Though speed was not mentioned in the instructions, all the *Ss* were set by self-instructions to learn their lists 'as soon as possible.' Though they did not know the purpose of the experiment, they did know that others from among their classmates were serving in the study, hence they were highly motivated to do "the best they could," as they did not wish to be among the least capable.

We also learned that the *Ss*, all of whom were inexperienced in memorizing letter-number pairs, passed rather quickly from less to more efficient methods of learning. That this was done is clearly shown by our results. For example; the Control *Ss* (Part I) required a mean of 4.9 ± 2.0 trials to learn the first series presented to them and a mean of 2.9 ± 1.2 trials (Part II) to learn the second series. The *memoria technica* developed during the learning of the first series brought a mean improvement of 2 trials in learning and a mean decrease of 0.8 in the *SD*.

At first, the *Ss* of both groups tended to repeat the stimulus-pairs, "to stamp them in memory by way of repetition," but they soon discovered, since the intervals between the successive exposures in the learning-series were so long that various methods could be tried, that the formation of mnemonic devices was much more efficient in accomplishing their task. Though some of the paired-associates were learned throughout the study by 'sheer memory,' *i.e.* by direct association, $a \rightarrow b$, the vast majority of them were learned by mnemonic devices, *i.e.* by indirect association, $a \rightarrow c \rightarrow b$.

The *Ss* also soon discovered that learning by way of mnemonic devices involved two tasks: first, the formation of the devices; and secondly, the retention of them. As noted above under the discussion of 'reversals,' the devices themselves were sometimes forgotten. When this occurred, the Control and Experimental *Ss* had different tasks. The Control *Ss* had to 'rebuild' the mnemonic device at the following learning-trials from the 'old' stimulus-pairs that were retained in the series; the Experimental *Ss*, to 'replace' them with new constructs from new stimulus-pairs. 'Rebuilding' from 'old' stimulus-pairs was, as our results indicate, more difficult than 'replacing' from a 'new' stimulus-pair. This is due, as we believe, to the fact that the memory-trace of the device forgotten interferes with the construction of a new device for the 'old' stimulus-pairs, whereas interference is not involved in replacing the forgotten device by a new one from a 'new' stimulus-pair. This interpretation stands, it should be noted, in disagreement with Rock's assumption mentioned above; namely, that the retention within the stimulus-series of cards once learned and then forgotten—*i.e.* the 'reversals'—favors the Control *Ss*. According to our results the inclusion of the cards failed has through interference the opposite effect.

The mnemonic devices of our *Ss* in both groups were idiosyncratic. Some were simple and spontaneous, others elaborate and bizarre. The simple and spontaneous were given immediately with the exposure of the stimulus-pair, or a fraction of a second thereafter. In either case, the balance of the 5-sec. interval was given over to rehearsing and establishing it by way of repetition. Examples of this type are:

A 5: "auto-$500"; *S* had recently purchased an automobile and had made a down payment of $500.
AA 1: "par excellence" reported by one *S*; "top quality," by another. (This number was assigned to these letters by a table of random numbers. It should probably have been changed but it was not. *Ss*' successes did not, however, pile up here as this obvious relationship, strange to report, was utilized by few *Ss*.
B 7: "brother, 7 yr. old"; (*S* had a brother of that age.)
C 20: "see 20/20 vision."
HH 31: S's initials and waist-measurement.
KK 13: "Klu Klux, unlucky."
JJ 23: "23 jumping jacks."
R 45: "revolver, 45 caliber."
RR 11: "railroad, parallel tracks."
WW 24: "World War II, lasted 4 years (for us)."
X 3: "X the third letter from the end of the alphabet."
Y 2: "young brother, 2 yr. old."

The elaborate and bizarre devices were 'worked out' during the exposure-interval and the 2-sec. pause. If they were not, then the pair was 'missed'; and the Control *Ss* got another chance to 'figure out' an adequate device during the next learning-trial and the Experimental *Ss*, another stimulus-pair 'to work upon.' Any kind of association between the stimulus-pair, elaborate or bizarre that tied the pair together, sufficed. As one *S* explained, he sought for "any kind of association—just anything—it didn't matter whether it was silly or not." Another reported that "the connection between the letters and numbers was often far-fetched, but that didn't matter." A third said, "any outside thing—anything at all—it doesn't matter how ridiculous, will tie the pair together." Examples of this type are:

A 5: "first letter of the alphabet and the first number counting by 5."
C 20: "third letter of the alphabet; counting by 10, 20 is one less than 30, the third 10."
CC 6: "phonetic, c-c-six."
DD 43: "rhyme and rhythm, d-d-forty-three."
E 32: "E is the fifth letter of the alphabet, and $3+2=5$.
J 28: "My first initial and my age four years from now."
K 4: "the letter and the number look something alike."
L 49: "L is the Roman numeral for 50; 49 is just one less."
N 46: "change the first vertical line of N into 4 and the second into 6."
Q 10: "two circles, the first of which is a Q, separated by a straight line."
SS 33: "Letters and numbers would be the same if the top halves of either were reversed. Letters and numbers look alike."
U 12: "Uranium and Carbon 12 test used in archeological findings; connected through radioactivity involved in both."

During the intervals between successive learning-trials and recall-tests, the *Ss*, in general, rehearsed, not the letters and numbers exposed to them, but the mnemonic constructs they had devised. They could, therefore, at the following learning-trials, concentrate their attention and effort on the construction of mnemonical devices for the 'unlearned' pairs. Rock's temporal intervals, as well as his exposure-materials (letter-numbers), lend themselves to mnemonical learning. If the intervals had been a little longer and the *Ss* had had a little more practice, the *Ss* of both groups would, as we believe, have learned the stimulus-series in a single trial. For example; our most able *S*, one in the Experimental Group, who learned the stimulus-series in two trials in Part I and in one trial in Part II, explained his performances as follows: "In high school, I was a spotter for the football team and I became highly practiced in associating numbers with players. If I had a name, or could make one up, I could remember the number. When I could think of a name to go with the letter or letters on the cards exposed in this experiment, I could remember the number." He had, as his results show, little trouble in "thinking up" names to go with the letters.

Now, what has this type of learning to do with the problems at hand? In studying the effect of repetition, no factors other than the presence or absence of repetition should be involved; and, furthermore, no more than one repetition of a stimulus-pair should be made at one learning-trial. In Rock's experiment, and our duplication of it, both of these conditions were violated. In the first place, learning was not effected by repeating the stimulus-materials but chiefly by the construction of mnemonic devices which were rehearsed an unknown number of times during the 'free' intervals in the procedure. Secondly, as noted above, the intervals between successive exposures of the stimulus-cards were so long (5 sec.) that, if learned by way of repetition, an indeterminate number of repetitions could be made. The number of learning-trials, therefore, bore no known relation to the number of times the stimulus-pairs were repeated.

In his classical studies on memory and learning, Ebbinghaus outlined the conditions under which the effect of repetition should be studied.[10] He presented his stimulus-materials (nonsense-syllables, which are not conducive to mnemonic learning) at a rate so rapid (150 per min.; 0.4 sec. per syllable) that not more than one repetition could be made at one exposure; that none could be rehearsed, and that "the invention of special associations of the mnemotechnic type" would be difficult if not impossible.[11] He, moreover, set himself, by self-instruction, against learning by mnemonic devices.

[10] Ebbinghaus, *op cit.*, 24-26.
[11] *Idem*, 25.

Rock's experimental conditions, on the contrary, favored their invention. Since mnemonic devices are not formed gradually by way of repetition, but, as in the learning of other relations, at a single trial or not at all,[12] we are of the opinion that *both* groups learned the paired-associates by the single-trial method: the Control *Ss* under the handicap of interferences aroused by the retention of the stimulus-pair that had been missed; and the Experimental *Ss* without that handicap, because the stimulus-pair failed had been removed from the series. Rock's and our corroborating results seemed, at this point in our study, to be nothing more nor less than artifacts of the method employed. To test these tentative conclusions, Experiment III was conducted.

Experiment III: Elimination of Mnemonic Aids

In the preceding experiments, the long intervals (3-sec. exposure and 2-sec. pause) between the successive exposures of the stimulus-cards permitted the "invention of special associations of the mnemotechnic type," against which Ebbinghaus especially warned. In the present experiment, we sought to eliminate these "inventions" by shortening the exposure-time and eliminating the inter-exposure pause.

Procedure. With two exceptions, the procedure described in Experiment I was used in this study. First, the exposure-time was reduced to 1 sec. and the pause between successive exposures was eliminated that the *Ss* would have no time to rehearse; and secondly, the number of stimulus-cards was reduced from 12 to 8 that the experiment could be completed in the time available before the beginning of final spring examinations.[13]

The *Ss* of both groups were presented, in haphazard order, the following stimulus-cards, which were chosen by chance from the stimulus-pool: **T-39**; **E-32**; **LL-38**; **JJ-23**; **PP-30**; **L-49**; **P-27**; and **Z-37**. In the experiments with the Control *Ss*, these cards were not changed, except in the order of their presentation—they were, as will be recalled, shuffled after every recall-test. The cards missed by the Experimental *Ss* were, however, withdrawn from the stimulus-series and replaced by new cards, chosen by chance, from the pool. The replacements were the same for all of these *Ss* that we could again determine whether the 'failures' clustered around certain stimulus-pairs.

Subjects. The *Ss* (30 in number, 28 undergraduate and 2 graduate students in psychology) were divided by chance into two groups of 15 *Ss* (4 women and 11 men), hereafter referred to as Control and Experimental. All of the *Ss* were volunteers. They were naïve in experiments of this type and served without knowledge of the problem.

Instructions. The instructions were also unchanged. We did not instruct the *Ss* to avoid the use of mnemonics, as Ebbinghaus did in his own case, because negative instructions often, especially with untrained *Ss*, prove positive. We were of the opinion, therefore, that the least said about mnemonic devices, the better.

---

[12] *Supra*, Footnote 5.
[13] Experiments I, II, and III were performed during the spring semester of 1957.

Results. The results of this experiment are given in Table III, which shows the mean and median number of learning-trials, the *SD*s of the means, and the ranges required by both groups of *S*s to reach our criterion of learning. As these data show, the performances of the Experimental *S*s, who learned by single-trial, were superior to those of the Control *S*s, who learned by way of repetition. The mean number of learning-trials required by the Control *S*s was 10.01; by the Experimental, 5.53. The difference (4.48) between these means, $t = 2.59$, is significant at the 2% level. The *SD*s of their means (3.06 and 1.77, respectively) and their medians (11.49 and 6.10, respectively) yield differences (1.19 and 5.39) which are also significant at the 2% level.

The mean number of stimulus-pairs reported correctly at every recall-test is also shown in Table III. At the first recall-test the Control *S*s had learned

TABLE III

MEAN AND MEDIAN NUMBER OF TRIALS REQUIRED FOR CRITERION, AND MEAN NUMBER OF STIMULUS-PAIRS LEARNED AT EVERY RECALL-TEST (Experiment III)

*S*s	*M*	*Med.*	*SD*	Range	Recall-test 1	2	3	4	5
Control	10.01	11.49	3.06	5–20	0.80	1.86	1.96	3.20	3.73
Experimental	5.53	6.10	1.77	3– 9	1.13	3.86	5.06	—	—

a mean of 0.80 cards, the Experimental, 1.13. This might, at first glance, be regarded as indicating that the better 'learners' were by chance assigned to the Experimental Group as conditions—even the stimulus-cards—were identical at the first learning-trial. The difference (0.33) between the means at the first recall-test is, however, statistically insignificant, hence we may assume that the *S*s of the two groups did not differ significantly in their ability to learn by the method of paired associates.

At the second recall-test in Experiment III, the mean number of cards correctly reported was 1.86 by the Control *S*s and 3.86 by the Experimental *S*s—a difference of 2.00 trials. At the third test, the means were 1.96 and 5.06, respectively—a difference of 3.10 trials. The differences between the means (2.00 and 3.10, respectively) are both statistically significant. The means at the later recall-tests could not be computed for the Experimental *S*s because it was the last test at which all of them reported—two of them reached criterion at this test. The means of the Control *S*s were, however, computed to the fifth recall-test as this test was the last at which all of them reported—one reached criterion at this point. The mean of the Control

Group at this test (3.73) was less, however, than that of the Experimental Group (3.86) at the second test. These results clearly show, in corroboration of Rock, that, under the conditions of this experiment, stimulus-pairs presented but once before learning are brought to criterion more rapidly—and significantly so—than pairs that are repeated until that level is attained.

Ss' comments. What the *S*s were doing during the exposure of the stimulus-materials and the intervals in the procedure was again an important matter, hence, at the conclusion of the experiment, they were asked to report upon the methods they used in learning the materials. They were, in particular, questioned regarding their use of mnemonic devices. Only four of the *S*s, one from the Control and three from the Experimental Group, reported that they had used these devices, but only when they arose spontaneously with the exposure of the stimulus-cards. None reported that he had invented them, as the *S*s in Experiment II did. Evidently the cards were exposed too rapidly to permit the construction of these devices. During the other intervals of the procedure, the *S*s did, however, rehearse at every opportunity as they were highly motivated and used every possible aid.

The effect of the mnemonical aids is clearly shown by a comparison of the *S*s performances in Part I of Experiment II, in which these devices were used freely, and in Experiment III, in which they, according to the reports of our *S*s, were rarely used. The mean number of trials required for criterion by the Control *S*s was 4.9 ± 2.0 in Experiment II and 10.01 ± 3.06 in Experiment III; and by the Experimental *S*s, 3.9 ± 1.1 and 5.53 ± 1.77, respectively. The differences between the means in these experiments are much greater than the figures show because the *S*s were required to learn 12 stimulus-pairs in Experiment II and only 8 in Experiment III. Eight pairs without mnemonics required approximately twice the number of learning-trials as 12 pairs with these aids. It would undoubtedly have been better had we used shorter exposures that mnemonical devices would have been entirely eliminated, and had we instructed the *S*s against rehearsing that the number of learning-trials would truly represent the number of times the stimulus-materials were repeated.

The results of Experiment III take us back to where we were before Experiment II was undertaken, but with this difference: We now know (1) that mnemonics are not responsible for Rock's results; and (2) that the task of the *S*s learning by single trial is *not* made easier by the elimination of the 'difficult' stimulus-cards "by the process of 'natural selection'," but rather that the task of learning by repetition is made *more* difficult, by way of associative interference, when the cards not learned are retained within the learning-series.

Our results in this series of experiments confirm Rock's. They justify us, as we believe, in concluding with him 'that associations are formed at a single trial and that repetition plays no role in their formation, as distinct from strengthening them, other than providing occasions when they may be formed at a single trial.'[14] Further evidence, with different learning-methods and materials and with animals, is highly desirable. The method of paired-associates and letter-number stimulus-material may be conducive to single-trial learning.

Supplementary Experiment

The following fall, 6 mo. after the completion of Experiment III, the *Ss*, without previous warning, were recalled and tested for retention. Of the *Ss*, six Control and four Experimental were available. We wished to determine which method, the repetitive or single-trial, yielded the greater retention.

The *Ss* of both groups were presented the stimulus-series that they had brought to criterion the previous spring. The apparatus and procedures were the same as used in Experiment III, with this difference: both groups learned by the repetitive method that neither group would have an advantage over the other; that the difference between their spring and fall performances would be entirely a matter of retention. The results are given in Table IV.

---

[14] In his discussion of multiple-item learning, Rock (*op cit.*, 1957, 192) points out that "a few seconds after any given pair is exposed during learning, and before the next pair is shown, . . . there is little doubt that most *Ss* would give the correct response," and asks whether in this instance, "an association was in fact formed"? He decided against it, and with his decision we are in complete agreement, but for very different reasons than those he gives. The ability to reproduce stimulus-material immediately after its exposure, or during a few seconds thereafter, is not a matter of memory, *i.e.* learning, but an example of a very different phenomenon; namely, the memory after-image (*Erinnerungsnachbild*) first observed and described by G. T. Fechner (*Elemente der Psychophysik*, 2, 1860, 494). It had better be called 'memory after-sensation' as it is a sensory phenomenon which persists for a few moments in 'memory' like a positive after-sensation and then is gone. Fechner describes the experience in the visual field only, but it also occurs, as C. C. Dimmick (The auditory memory after-image, this JOURNAL, 34, 1923, 1-12) shows, in the auditory field. She reviews the literature on the phenomenon and describes it and the conditions under which it occurs. E. B. Titchener (*A Beginner's Psychology*, 1915, 74) illustrates the auditory memory after-image by reference to taking dictation or notes in the classroom. He says, "as you write the words last spoken, the speaker's voice still rings in your ears; the sound hangs for a few seconds, as if arrested, and your pen is guided by the mental echo." An illustration of the visual type is often experienced by the senior author of this paper. After looking up a telephone number in the directory, he frequently has to look it up again if he is interrupted, even briefly, before he dials it. The memory after-image, unless the letters and numbers are repeated aloud—unless auditorially and kinesthetically supplemented—are of short duration. Indeed, so short that the call numbers, unless supplemented, are frequently 'forgotten' before the latter ones are dialed. He has learned, through experience, to repeat the call number aloud several times before closing the directory. Perhaps all learning is of this nature: A memory after-image of the stimulus-materials, which follows the perception, is lost in seconds unless it is supplemented by mnemonical devices or repetition, or primacy, recency, vividness, etc., or by factors as yet undiscovered, which, when operating alone (*i.e.* when mnemonics, repetition, etc., are eliminated) are responsible for single-trial learning.

Results. A comparison of Tables III and IV reveals that the *Ss* available during the fall were fair samples of their respective groups. Their mean number of learning-trials required for criterion in the spring experiment was very similar to that of their groups, viz. the mean of the Control Group was 10.01 ± 3.06, of the Control Sample, 10.50 ± 3.80; of the Experimental Group, 5.53 ± 1.77, of the Experimental Sample, 5.25 ± 2.79.

The Control and Experimental Samples relearned, in the fall, the series they had mastered the previous spring with mean savings of 1.67 and 1.25 trials, respectively. Four of the *Ss* in the Control Sample relearned in fewer trials than in the spring, one in the same number, and one required more learning trials (11 against 10). Three of the *Ss* in the Experimental Sample, though using the less efficient method

TABLE IV

MEAN NUMBER OF TRIALS REQUIRED FOR CRITERION IN THE SPRING AND FALL EXPERIMENTS

(Supplementary Experiment)

Control			Experimental		
Ss	Spring*	Fall*	*Ss*	Spring†	Fall*
1	20	20	1	3	2
2	10	11	2	8	6
3	5	4	3	3	3
4	9	3	4	7	5
5	10	7			
6	9	8			
Mean	10.50	8.83	Mean	5.25	4.00
SD	5.09	6.19	*SD*	2.79	1.79

* Learned by the repetitive method.
† Learned by single-trial method

of learning (the repetitive in place of the single-trial method) did better than in the spring and one showed no improvement. The one showing no improvement had, however, less room for improvement as he learned and relearned in three trials.

Though the data are too few in number to warrant statistical treatment, a *t*-test of the difference (4.43) between the means of the Control and Experimental Samples was computed. It was 1.82, which is insignificant at the 10% level. The experiment does, however, reveal that the *Ss* from the Control Sample relearned with a saving of 15.9% and those from the Experimental Sample relearned with a saving of 23.8%. The difference (7.9%), though not highly significant, is in the direction of the Experimental *Ss* whose savings, as will be recalled, with a less efficient method of learning (the repetitive method) were computed from the results obtained by a more efficient method (single-trial).

The results of the supplementary experiment seem, therefore, to indicate that retention of the paired-associates, in addition to the rate at which the associations are formed, is superior in single-trial to repetitive learning.[18] This may possibly be due to the theory, discussed in Experiment II, that fewer associative interferences are aroused in learning by single trial than by repetition.

[18] Wogan and Waters, in their study mentioned in Footnote 7, confirm this conclusion.

Summary

Three experiments were conducted to test Rock's results that paired-associates presented but once before learning are learned as rapidly as those repeated until learned; from which he drew the conclusion that "repetition plays no role in the formation . . . of associations, other than that of providing the occasion for new ones to be formed, at a single trial." Experiment I, a replication of Rock's, corroborated his results.

Experiment II, undertaken to determine whether these results are due to the elimination of idiosyncratically difficult cards from the stimulus-series, indicated that the elimination of 'difficult' cards did not affect the task in single-trial learning but that the inclusion of those cards in repetitive learning did make the task more difficult by way of interference. This experiment also showed that the long exposure-time, the long and numerous 'free' intervals between successive phases in the procedure which lend themselves to rehearsing, the stimulus-materials, and the method of paired-associates are conducive to mnemonical learning to which the single-trial and repetitive learners soon resorted. Since these mnemonic devices are constructed at one trial or not at all, the similarity of the results of the two groups which were very like, is not surprising. Both groups learned by the same, the single-trial, method. The results of Experiment I seem, therefore, to be artifacts of the method.

Experiment III, undertaken to determine what effect the reduction of the exposure-time to 1 sec., and the elimination and shortening of the intervals between successive phases in the procedure have, yielded surprising results. They showed, in corroboration with Rock's and our results in Experiment I, that single-trial learning was significantly superior to repetitive learning when mnemonical devices are eliminated or greatly reduced.

After giving suggestions for further studies, the results of a Supplementary Experiment on retention were reported. The *S*s relearned the stimulus-series they had learned 6 mo. before. The savings in relearning indicate that retention, as well as rate of learning, is superior in single-trial to repetitive learning—which placed us back where we were before this series of experiments was undertaken. We duplicate Rock's results but are still unwilling to accept his conclusions. They seem to be inescapable, but we still wish more evidence.

ESTIMATED PROBABILITY OF SUCCESS FOR A FIXED GOAL

By JAMES C. DIGGORY, EUGENE J. RILEY, and RUTH BLUMENFELD,
University of Pennsylvania

Since 1930, the antecedents and consequences of success and failure have been studied extensively by means of the 'level-of-aspiration' paradigm developed by Hoppe.[1] It was early discovered that the level of performance attempted (*i.e.* the goal or level of aspiration)[2] is not generally the same for each trial in a series, and the average of these goals varies from one subject (*S*) to another. The bulk of the literature on level of aspiration represents attempts to understand the sources of these variations: the problem of how goals are set.[3] Generally, if an actual performance equals or exceeds the level intended, the goal or intended level of the next performance is higher, and if the performance falls below its intended level, the goal of the next performance is lower. Thus, as performance varies in a series of trials, the goals also vary. This variation of goal from one performance to another is, however, a consequence of the methods for study in the level of aspiration which have not led to consideration of the case where *S* attempts, in a series of performances, to achieve a *fixed* goal. This paper presents a series of experiments dealing with some aspects of the latter case.

There are situations in which people strive for goals which they themselves cannot change because the goals are standards of performance established and maintained by someone else, *e.g.* attempts to attain the minimal qualifications for entry into schools, trades, professions, and social groups, or for advancement in rank or status. In these cases, the discrepancy between goal and actual performance is always negative until

* Received for publication February 28, 1958. The completion of this study was made possible by a grant-in-aid from the University of Pennsylvania Committee on the Advancement of Research.

[1] Ferdinand Hoppe, Erfolg and Misserfolg, *Psychol. Forsch.*, 14, 1931, 1-62.

[2] The term "level of aspiration" has been used in many other senses, including the performance "expected," "hoped for," or "feared" on the next trial. See, J. C. Diggory, Responses to experimentally induced failure, this JOURNAL, 62, 1949, 48-61, and M. G. Preston and J. A. Bayton, Differential effects of a social variable upon three levels of aspiration, *J. exp. Psychol.*, 29, 1941, 351-369. We here restrict it to the original definitions of Hoppe, *op. cit.*, and J. D. Frank, Level of aspiration test, in H. A. Murray, *et al., Explorations in Personality*, 1938, 461-471.

[3] Summarized by Kurt Lewin, Tamara Dembo, Leon Festinger, and P. S. Sears, Level of aspiration, in J. McV. Hunt (ed.) *Personality and the Behavior Disorders*, 1, 1944, 333-378.

the goal is achieved. Thus, a person has no clear grounds for assurance that he will reach the goal. It is, nevertheless, a common observation that of two persons, whose performance-levels are objectively equal, one decides to continue striving for a goal while the other abandons it. Our general problem is to suggest a plausible basis for these different decisions.

We assume that a person's perseverance toward an objective, or his withdrawal from its pursuit, are determined by his perception of *the probability that the discrepancy between goal and actual performance can be reduced to zero.* We shall, hereafter, refer to this as the probability of success, P(s), since reduction of the discrepancy to zero defines success. When goals are approached gradually, through a series of performances, the striver may relate each performance to some sub-goal which he will try for on the next attempt, *i.e.* a level of aspiration in the usual sense. In the conditions we are considering, these do not represent the success for which the individual is striving although their successive achievement informs him of his distance from the goal and his rate of progress toward it. Nor is improvement in performance, by itself, a guarantee of reaching the goal, because there is usually a 'deadline,' imposed by the performer himself or by other persons or conditions, beyond which no further opportunities for pursuit of the goal will occur. Thus, for P(s) to be high enough to warrant a person's continuing, his distance from the goal must be small enough, and his rate of improvement large enough, to support the expectation that he will achieve the goal before his opportunities are exhausted. For many rates of progress, P(s) should be related directly to the number of opportunities remaining, so that as one gets closer to the deadline, his estimate of P(s) ought to decline, although his performance is actually improving.

Our interest in P(s) was derived from some theoretical considerations about the manner in which people evaluate themselves. If one puts aside for the moment such general notions as 'value of the total self' or 'value as a person,' it is possible to consider whether people evaluate themselves as good, bad, or indifferent instruments for the pursuit of specific goals, such as winning chess games, analyzing a body of literature, carving wooden figures, or solving differential equations.[4] In tasks like these,

[4] For other suggestions about this utilitarian view of self-evaluation see, for example, Gardner Murphy, L. B. Murphy, and Theodore Newcomb, *Experimental Social Psychology* (Rev. ed.), 1937, 209; Gardner Murphy, *Personality*, 1947, 523; Leon Litwinsky, Toward the reinstatement of the concept of the self, *Brit. J. Psychol.*, 1951, 42, 246 f.; C. M. Solley and Ross Stagner, Effects of magnitude of temporal barriers, type of goal, and perception of self, *J. exp. Psychol.*, 51, 1956,

proficiency may improve to such a level that it serves as a means of acquiring status, power, or livelihood. Many a person who pursues these objectives has a relatively high estimate of his own instrumental value while still in quest of his goal, *i.e.* before he actually achieves it. We suppose that this is because he estimates P(s) to be high. Conversely, we believe that those who, in spite of recognized past accomplishments, hold themselves in relatively low esteem, do so because they cannot progress any farther, *i.e.* P(s) for further achievement is low. Thus we have attempted to understand the performance-conditions which lead to high or low estimates of P(s) in the belief that we will thereby increase our understanding of how people evaluate themselves as instruments for the achievement of specific goals. Following the considerations stated above, we set up conditions in which the *S*s were given several opportunities to achieve a fixed standard of performance, with systematic variation of characteristics of the performance-curves, distances from the goal and the deadline, and *S*'s motivation.

PROCEDURE

All *S*s were men, undergraduate students, recruited indiscriminately from the College of Arts and Sciences, the Wharton School of Finance and Commerce, the School of Fine Arts, and the Engineering Schools, at the University of Pennsylvania. All were given individual ½-hr. appointments, with the exception of those in Experiment III for whom the standard appointment was 1 hr. Four *S*s were assigned randomly to each condition in all of the experiments.

In every experiment, *S* made successive estimates of P(s) while being completely informed about his performance, goal, and deadline. These items of information were presented in a visual display using pegs inserted in a perforated Masonite panel, shown in Fig. 1. The horizontal string marks the goal. The vertical string was used to mark the deadlines. The procedures for the manipulation of the independent variables are the same for all experiments.

(1) Characteristic of the performance-curve. The apparent characteristics of the curve describing performance were controlled by inducing *S* to adopt the goal of sorting 40 cards into 10 suits within 25 sec. on at least one trial. The experimenter (*E*) manipulated a stopwatch, pretending to keep time, but actually stopping *S* after he had correctly sorted a predetermined number of cards for each trial. The performance-curves used are shown in Figs. 2 and 3. After every trial, *S*'s score, for that trial, was indicated by inserting a golf tee into the appropriate hole in the Masonite panel.

(2) Motivation. Motivation was manipulated by representing to some *S*s that the card-sorting task was actually a test of their basic ability to profit by formal training in the physical sciences. The remaining *S*s did exactly the same task, but it was

---

69; and Ezra Stotland, Stanley Thorley, Edwin Thomas, A. R. Cohen, and Alvin Zander, The effects of group expectations and self-esteem upon self-evaluation, *J. abnorm. soc. Psychol.*, 54, 1957, 55-63.

not represented as a test—only as a task in which we were interested. The *Ss* who took the 'test' volunteered to do so after receiving the following (false) announcement.

BASIC SCIENCE ABILITY SURVEY
U. of P. Project 372

The Psychology Department is coöperating in a nationwide survey designed to discover people whose basic abilities would make it worthwhile to give them training in the physical sciences in the event of a national emergency. The funds for this work come from private sources but the results of the study will be made available to Federal and State governments.

It is strongly suspected that many people who might contribute to the advancement of the physical sciences are not planning scientific careers nor even taking any training in these sciences. We are asking for volunteers to take a simple preliminary test. The outcome of the test will determine whether or not a person is eligible to be placed on the roster of those for whom such training might be profitable.

FIG. 1. PERFORMANCE DISPLAY BOARD
The curve shown corresponds to Condition 3 (Fig. 2) and to Curve A (Fig. 3).

We recognize that if a person has no interest in scientific work he is unlikely to do well in it even if he has the necessary basic abilities. Thus, as a preliminary measure of interest, we are asking only for volunteers for this test.

Even if you should pass the test, your name will not be placed on the list unless you authorize it. Moreover, the list will be reviewed continuously in the future so that those whose interests change can be eliminated. Thus, the list can be kept up to date.

James C. Diggory, Research Supervisor
Eugene J. Riley and Ruth Blumenfeld, Research Specialists

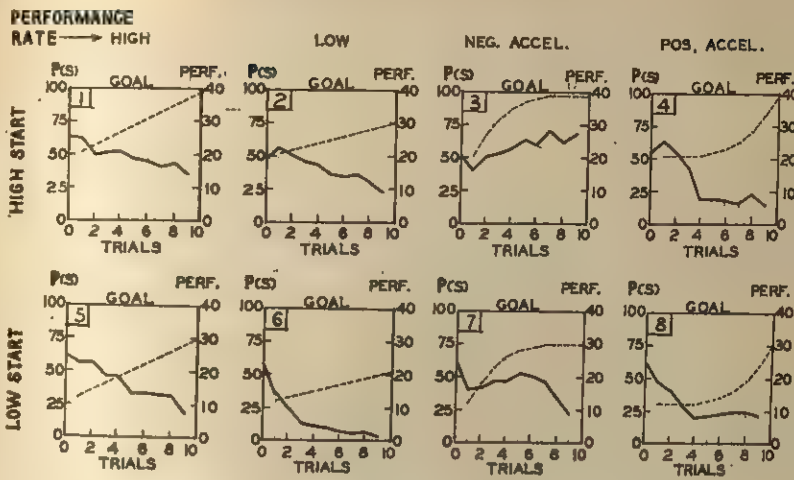


FIG. 2. ESTIMATES OF PROBABILITY OF SUCCESS AND PERFORMANCES REPORTED TO THE *S*s

Solid lines, estimate of P(s); broken lines, performance.

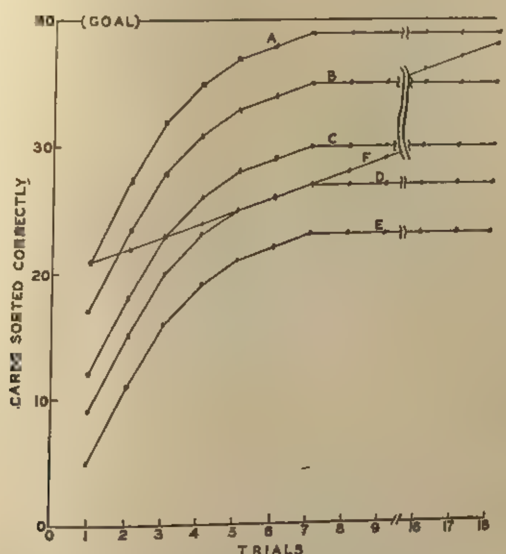


FIG. 3. TREATMENTS UTILIZED IN EXPERIMENTS II AND IV

The curves show the performance reported to *S*. The letters refer to treatments of the data.

Those who responded to this appeal are hereafter called highly motivated *Ss* (Hi-Mot). When a highly motivated *S* came to his appointment, he immediately filled out a form authorizing the placing of his name on the 'list.' This form was to be countersigned later by *E*, the 'Test administrator,' if *S* passed the test.

The low motivated *Ss* (Lo-Mot) were simply told that we would like to see them try to sort 40 cards in 25 sec. They were recruited by inviting them to participate in an interesting psychological experiment which was neither embarrassing nor dangerous. They were told that they would receive a full written account of the experiment as soon as the data were analyzed. This was done in every case, even for the highly motivated *Ss*. Furthermore, all the *Ss*, who were students in psychology, were rewarded for their participation in the experiment by relieving them of the necessity of writing one of their term reports. The experiment was explained to the *Ss* only *after* they had completed it. None of them knew about the real nature of the experiment during its course.

It should be noted that the expressions 'high motivation' and 'low motivation' are used rather loosely here, since we cannot demonstrate that the motive of wanting to become a scientist is stronger than to avoid writing a report or to satisfy curiosity. It can be plausibly argued, however, that wanting to become a scientist is more relevant to an individual's self-esteem than either of the other two objectives. Furthermore, the rewards offered to successful passers of the 'test' were much more valuable than those offered to the other *Ss*, and finally, as the sequel will show, the highly motivated *Ss* behaved in a systematically different way from the low motivated *Ss*.

(3) Clarity of the deadline. The deadline to which the *S* worked was manipulated to create two conditions: vague and clear. After the requirements of the task had been described, the *Ss* in the 'clear' condition were told, "You will have several opportunities to do so, but we will stop if you do it once. Since our time is limited, we will have to stop sometime, but you will have more than 10 trials. After the tenth trial I will tell you how many more trials you will have." Also, in the 'vague' condition, the vertical string used to mark the deadline (see Fig. 1) was absent at the beginning, and was put in place only when a deadline was actually announced after Trial 10.

The task was card-sorting, and the following instructions were read to each *S*, the phrases referring to a 'test' being omitted for the low motivated *Ss*.

The basic task [in this test] is to sort cards according to a pattern I will set up. Each card has a pattern on its face and there are 10 different patterns. You will be given a well-shuffled pack of 70 cards which you must hold face down in one hand. When I say go, take the top card of the pack, turn it over and place it on top of the card on the table which has the same pattern on it. Do the same with the next card, and so on, until I say stop. [To pass the test] you must correctly sort at least 40 cards in 25 sec.

This [test] task may seem very simple, but actually several basic processes must be well-coördinated to do a good job: *memory* for the forms and their locations, rapid *perception* of the forms as they appear, and *manual dexterity* in handling and placing the cards.

Notice that on the board there is a string stretched at level 40 to indicate [the passing score] the score you are to try for. There is another string here [point] just after 10, to remind you that you will have 10 trials [used only for clear deadline conditions, see above]. I will show you your score after each trial by putting a peg in one of these holes.

I will count only cards that are correctly placed, so if you sort 35 cards, but two are wrongly placed, your score is 33. Do you have any questions? I am going to ask *you* a question.

At this point, *S* was given a copy of the following scale on a slip of paper and was asked to fill in trial number and score (Trial 0, score × *before* his first trial) and answer the question.

Trial No. ———— Score ————

What, do you think, are your chances of having your name placed on the list by scoring 40 or more on this test? [Hi-Mot *Ss* only.] What, do you think, are your chances of sorting 40 cards in 25 sec. on at least one of the remaining trials? [Lo-Mot *Ss* only.] Mark the scale anywhere to show your estimate.

0	10	20	30	40	50	60	70	80	90	100
Can't possibly do it		Less than even chances I can			Chances are even that I can		More than even chance I can			Can certainly do it

S repeated this estimation of P(s) after each successive trial, except the last when the outcome was known with certainty and questions about P(s) were irrelevant. Also, in Experiment IV, estimates of P(s) were taken only for Trials 0, 5, and 9.

Experiment I

All *Ss* were recruited for the highly motivated condition and their scores were manipulated to produce four different performance-curves, each starting at two different levels. These define the eight experimental conditions and are represented by the broken lines in Fig. 2. Each condition is also numbered from 1-8.

This figure also presents, for each condition, the average P(s) estimates of four *Ss* for each trial. The expectation that nearing the deadline would lower P(s) is clearly confirmed for all but Condition 3.

Analysis of the trends in the P(s) by Alexander's method shows that the over-all slope differs significantly from zero, the slopes under the eight conditions differ from one another, as do the slopes for the individual *Ss*.[5] There are differences between individual means due to the fact that the *Ss* started their estimates of P(s) at different levels and there is some tendency for subsequent estimates to be correlated with the initial level. The differences among group means cannot be attributed to differences in mean group starting-level since analysis of variance of the initial estimates of P(s) (Trial 0) reveals no significant differences among conditions. Thus, the effects noted between groups are due to the experimental

[5] H. W. Alexander, A general test for trend, *Psychol. Bull.*, 43, 1946, 533-557. Our confidence-level is $\alpha < 0.05$. Analysis of variance of the estimates at Trial 0 reveals no significant initial differences among *Ss* assigned to the various conditions in any of the experiments.

treatments. The differences between group means are a consequence of different rates of decline in P(s).

Differences in initial imputed *performance* produce a decrease in P(s) that is significantly greater for the low than for the high starting group, as can be seen in Fig. 2.[6] A constant and positive rate of improvement in performance (Conditions 1, 2, 5, and 6), is generally associated with a steady decline in P(s), though the higher rates of improvement produce smaller declines in P(s), and the effect of rate depends on mean distance from the goal. In fact, the four conditions are ordered with respect to the slopes of their P(s)-curves as follows.

Performance		Condition	Slope of P(s)-curve
Start	Rate		
High	High	1	−2.64
	Low	2	−3.23
Low	High	5	−4.77
	Low	6	−5.00

Generally, zero rate of improvement also produced a sharp decline in P(s) (Conditions 4, 7, and 8), the single exception being Condition 3, where the plateau is only one point below the goal. It is interesting to note that in Conditions 4 and 8, where the plateau occurs in the first four trials, the subsequent rapid improvement in performance is not accompanied by any significant recovery in P(s).

Experiment II

In this experiment motivation and mean distance from the goal were the independent variables. Motivation was varied by the method outlined above. Four *S*s subjected to each of the two motivation-conditions were assigned randomly to each of the treatments indicated by Curves A, B, C, D, and E of Fig. 3; a total of 40 *S*s. It should be recalled that these curves represent the reports *E* gave *S* regarding his ostensible performance during a fictitiously constant 25-sec. period. The experiment was performed twice for reasons explained below, so there are 80 reports altogether. Fig. 4 shows how mean distance from the goal affects the mean estimate of P(s).

[6] Mann-Whitney U-Test, $\alpha < 0.05$.

The means and slopes of the curves differ significantly and the slopes are non-zero and non-linear.

The effects of motivation on P(s) are shown in Fig. 5. When these conditions are compared, the means of the two groups are significantly

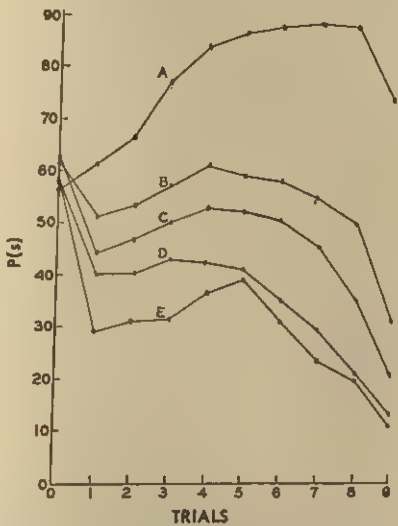


FIG. 4. P(s)-CURVES FOR DIFFERENT DISTANCES BETWEEN PERFORMANCE AND GOAL

The letters refer to the corresponding performance-curves in Fig. 3. Every point is the mean of 16 *Ss*.

different; the over-all slope differs significantly from zero and from a linear trend-line, but there is no difference between the group-slopes.

The data presented above are actually a combination of two identical experiments, run about 2 mo. apart. We made the replication because, in the first experiment, plotting mean estimates of P(s) per *S* per trial for each of the 10 conditions in the manner of Fig. 6 revealed an interaction between motivation and distance from the goal: the amount by which Hi-Mot exceeded Lo-Mot *increased* as distance from the goal increased. The purpose of the replication was to see whether this was a fact, but the interaction does not appear in the second set of data nor when the two sets of data are combined as shown in Fig. 6.

EXPERIMENT III

The effects of performance (Curves A and F, Fig. 3) and 'nearness' and 'clarity' of the deadline were dealt with in Experiment III. All *Ss* were highly motivated and received the instructions giving them a vague

deadline. Experimental appointments were made for 1 hr. to provide sufficient time for an extended series of trials after the tenth. Distance from the deadline after Trial 10 was varied by announcing that *S* had either 1, 2, 4, or 8 trials remaining. Thus, there are eight experimental condi-

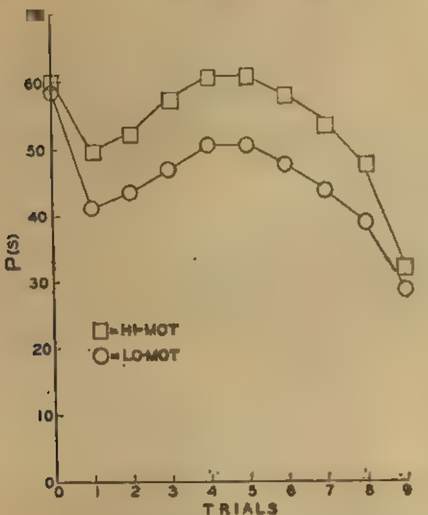


FIG. 5.

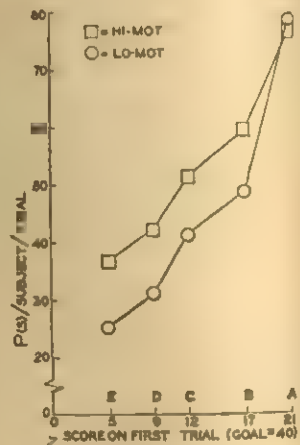


FIG. 6.

FIG. 5. P(S)-CURVES FOR HIGH AND LOW MOTIVATION
Every point is the mean of 40 *Ss*.

FIG. 6. P(S) AS A FUNCTION OF MOTIVATION AND DISTANCE FROM GOAL
Every point is mean of 8 *Ss* × 10 trials. Letters refer to performance-curves in Fig. 3.

tions, two for characteristics of the performance-curve, and four for distance from the deadline, with four *Ss* assigned randomly to each condition.

Fig. 7 shows estimates of P(s) over Trails 0–9 in conditions where the deadline was vague. At Trial 10, the deadline was announced to be 1, 2, 4, or 8 trials more. This announcement was made *after S* had completed Trial 10 but *before* he knew his score for that trial, thus the estimate of P(s) for Trial 10 belongs to the clear deadline-condition. This is the point in the experiment at which sudden revelation of nearness to the deadline might be expected to be effective, so we analyzed the changes in the estimates of P(s) between Trials 9 and 10. These mean changes, associated with their respective distances from the announced deadline, are as follows.

Number of trials after Trial 10	1	2	4	8
Mean change of P(s) Trials 9–10	−5.25	−3.00	0.00	+1.13

Obviously the correlation between distance from the deadline and change in P(s) between Trials 9–10 is high. Another sense in which nearness to the deadline has an effect is shown by the course of the estimates of P(s) following Trial 10. Once the deadline has been established, every successive trial reduces the number of remaining opportunities, or to put it another way, brings *S* closer to the deadline. As Fig. 7 indicates a

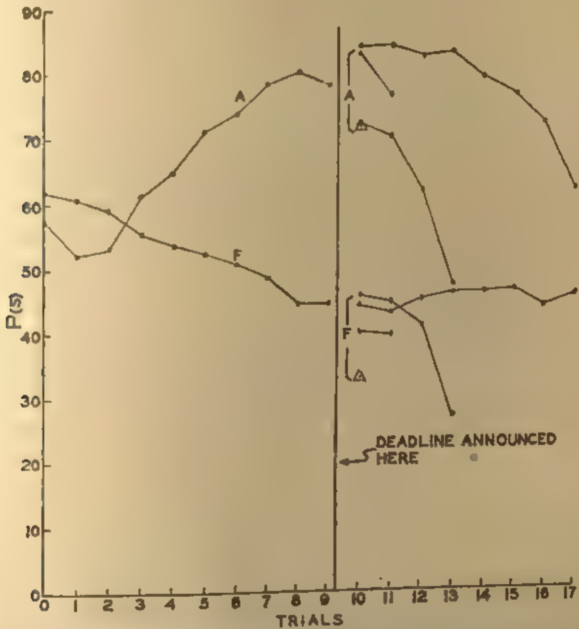


FIG. 7. P(s) AS A FUNCTION OF VARIOUS DISTANCES FROM DEADLINE AND TYPE OF PERFORMANCE-CURVE (A AND F, FIG. 3)

Every point for Trials 1–9 is the mean of 16 *Ss*; after Trial 9, every point is the mean of 4 *Ss*.

general decline in P(s) occurs in this part of the data, with the exception that in Curve F, with 8 trials remaining, the decline is halted after Trial 10, no doubt because of the continued steady approach to the goal.

EXPERIMENT IV

In Experiment IV, we used two levels of motivation (Hi-Mot vs. Lo-Mot), two levels of distance from the goal Curves A and C of Fig. 3), and two levels of clarity of the deadline (vague and clear); all varied according to the statements given in the procedures' section. Estimates of

P(s) were made only at Trials 0, 5, and 9. Fig. 8 A shows the effects of motivation; Fig. 8 B, the effects of vagueness of deadline; and Fig. 8 C, the effects of distance from the goal. Clarity of deadline and distance from the goal have significant effects on both means and slopes.

The broken curves in Fig. 8 are drawn from the previously reported experiments under conditions comparable with those of the present experi-

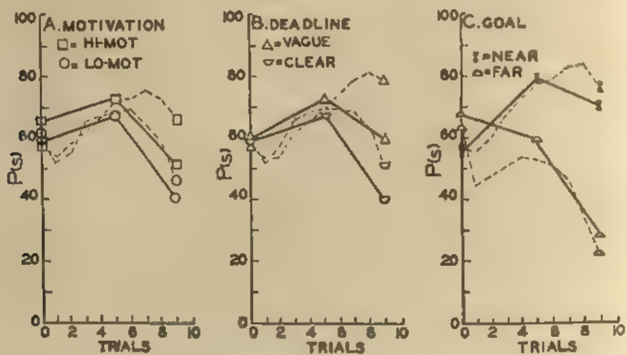


FIG. 8. P(S) AS A FUNCTION OF MOTIVATION, CLARITY OF DEADLINE, AND DISTANCE FROM THE GOAL

Solid lines are for Experiment IV, broken lines are for comparable conditions from other experiments.

ment. They are presented to facilitate comparison of the estimates of P(s) made at all Trials 0–9 with those collected only at Trials 0, 5, and 9. There is one respect in which the condition of vagueness of deadline is not strictly comparable between Experiment III and this one. In Experiment III, all the *S*s were recruited for a 1-hr. session, whereas in the present experiment ½-hr. sessions were provided. Thus, in the previous experiment *S* could realistically expect an extended number of trials, but in the present one he could hope for only one or two trials after the tenth, because it was obvious, by the time this trial was completed, there was little time left for additional attempts. Nevertheless, the differences between the partially sampled and the completely sampled P(s)-curves are very small and are maximal only at Trial 9, where the difference noted above would be expected to have its greatest effect. Indeed, Alexander's trend-test applied to the difference between Trials 0, 5, and 9 of the present experiment and those of the earlier experiments indicates no significant differences in means or slopes.

In the present experiment, the motivational variable did not have a

significant effect. This is probably due to the small number of *S*s, because the differences due to motivation are in the same direction and about the same magnitude as those in Experiment II. The motivational variable is not, however, completely without effect because the difference between motivation conditions is significant when the deadline is vague, though not when it is clear.

The fact that the results of the present experiment agree with those from the earlier controlling experiments means, at least within the limits of our experimental conditions, that estimates of P(s) are more dependent upon motivation, nearness to the deadline, and distance from the goal, than they are upon each other. That is, if serial effects occur, they are not important enough to concern us at the moment. This is of some technical interest because it provides an opportunity for using 'blank' trials to collect other types of data; for example, statements about the level of aspiration.

In addressing ourselves to the effects of our experimental variables, we give little attention to the individual differences among *S*s in means and slopes of the P(s)-curves.[7] Although such individual differences were found in every experiment, they were not great enough to swamp the effects of the experimental conditions. In the context of our interest in self-evaluation these differences are important and our failure to report on them more completely is only temporary, pending the adoption or invention of special procedures for their study. Once we learn to control or predict individual differences in estimates of P(s), our knowledge will be increased and our experiments can be done with more homogeneous groups of *S*s and thus become more precise.

Discussion

The experiments reported support our contention that as a person strives for a fixed goal in a series of attempts, his estimate of the probability of his success, P(s), depends directly on his rate of improvement, his nearness to the goal, and his distance from the deadline. The finding that P(s) also varies with the presumed importance of the goal to *S* was a surprise because we had no way or predicting it. We can now see that it is related to the general question of the relation between the attractiveness (valence) of a goal and the expectation that one will achieve it. Perhaps the apparent attractiveness of a goal leads to confounding of judgments of its real value and its distance, analogously to the way in which the apparent size of an object in vision leads to confounding of judgments

[7] Differences among these slopes are significant ($\alpha < 0.05$) by Alexander's test.

of its real size and its distance. This and related questions cannot be answered until we have some way of getting *independent* measures of the value of a goal and the expectation of achieving it.

An equally important problem is the relation of P(s) to the level of aspiration. Atkinson recently presented a model for predicting (a) choices among tasks of varying difficulty and (b) the intensity of behavior (motivation) directed to a task of known or stated level of difficulty when *S* has no choice but to attempt it. He uses the term "probability of success" as functionally equivalent to the level of aspiration, and shows that many classic findings on the latter are consistent with predictions achieved with his model. If we had asked *S*, after each trial, the score he would try to achieve on the next trial, the resulting level of aspiration would have been either roughly parallel to the performance-curves or constant at the level of the fixed goal. Since our P(s)-curves do neither of these, we conclude that P(s), as we have used it here, is not the same thing as level of aspiration.[8]

Another area of research to which our findings may be related is that of group-processes. Is the function of uniformity in social groups the same regardless of the rate at which consensus is achieved or the number of available opportunities to achieve it (deadline conditions)?[9] Thus, if a need for self-evaluation produces a demand for unanimity of opinion among the members of a group, as Festinger says,[10] then it seems to us that the goal of unanimity may be approached at different rates and distances from a deadline, so that P(s) will be high or low with important consequences for the composition of the group—its cohesiveness and its status as a reference group.

As a final consideration we refer to our suggestion in the introduction that a person's estimate of his own P(s) for a given goal determines his evaluation of himself as an instrument for the pursuit of that goal. We must add to this assumption the provision that *S*'s approach to the goal depends only on his own efforts and not on luck or the kindness or maleficence of other people. For further discussion, we assume this link

---

[8] J. W. Atkinson, Motivational determinants of risk-taking behavior, *Psychol. Rev.*, 64, 1957, 359-372. We agree with Atkinson that the two are related, but we disagree with his view that the correspondence is one-to-one.

[9] See, for example, Stanley Schacter, Deviation, rejection, and communication, *J. abnorm. soc. Psychol.*, 46, 1951, 190-207. His analysis of the treatment accorded to the 'slider' is relevant to our view, but neither the rate nor the deadline for 'sliding' was used as an experimental variable.

[10] Festinger, A theory of social comparison processes, *Hum. Rel.*, 7, 1954, 117-140; Motivations leading to social behavior, in M. R. Jones (Ed.) *Nebraska Symposium on Motivation*, 1954, 191-219.

between P(s) and self-evaluation to be a fact. Now consider a highly motivated person who is pursuing an important goal. His P(s) will, other things permitting, be high (Experiment II). Suppose further, that he makes progress toward his goal, but passes from youth to middle age in the process. It is by no means certain that his progress will continue to support high P(s), for now he is approaching a deadline (arbitrary retirement, senescence, or death) beyond which his striving cannot continue. Granted these assumptions, and our data, his self-evaluation is likely to decline.

Consider another person striving for an important goal with "his whole life before him," and suppose that he himself has imposed a deadline which is, therefore, not inevitable, but a matter of choice. If he will not move the deadline, his P(s) may decline and his self-evaluation with it. For how many people do some such conditions produce aimlessness, *ennui*, neurosis, or worse? This and related questions will guide our future research.

Summary

In attempting to reach a fixed goal within a limited number of trials, with full knowledge of their rate of progress and distance from the goal, *S*s varied their estimates of their probability of success, P(s), with rate of progress toward the goal, proximity to the goal, and distance from the deadline. It is argued that P(s), as it applied to this situation, is not the same as the level of aspiration. The relevance of the findings for group-processes is discussed. It is also argued that one's P(s) for a given goal determines his evaluation of himself as an instrument for achieving that goal.

RESPONSE TO CHANGING PATTERNS OF EVENTS

By Jacqueline J. Goodnow, Irvin Rubinstein, and Ardie Lubin,
Walter Reed Army Institute of Research

What is the effect of prior experience in a situation where the *S* faces a change in the probability with which a stimulus-event occurs or a response is reinforced? This is the type of change that occurs in extinction after various ratios of reinforcement, and also in studies of learning-sets. A number of psychologists have thought that neither a standard S–R approach nor an approach in terms of general expectations adequately describes the data from such experiments and have proposed instead some redefinition of the critical stimuli or the responses.[1] One recurring proposal is to consider stimuli in terms of the patterns or sequences they form and to regard probability-shifts as a problem of discriminating between patterns.[2] The difficulty with this proposal lies in pinpointing the critical patterns or sequences. The present study explores the value of defining them in terms of length of run—the number of identical stimulus-events occurring in a row before response on any particular trial.

Suppose *S* is required to predict on each of a series of trials whether Event *a* or Event *b* will occur. These events, *a* and *b*, are the stimulus-events. For the first 100 or so trials, Events *a* and *b* occur equally often (50:50). Then Event *a* starts appearing on 75% of the trials (75:25). This is a change in the over-all probability with which Event *a* occurs and,

* Received for publication May 9, 1958. For a generous supply of data and comment, we are indebted to Dr. Allen Parducci.

[1] David Friedes, Goal-box cues and pattern of reinforcement, *J. exp. Psychol.*, 53, 1957, 361-371; J. J. Goodnow and T. F. Pettigrew, Effect of prior pattern of experience upon strategies and learning sets, *ibid.*, 49, 1955, 381-389; D. A. Grant, H. W. Hake and J. P. Hornseth, Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement, *ibid.*, 42, 1951, 1-5; F. A. Logan, E. M. Beier, and W. D. Kinkaid, Extinction following partial and varied reinforcement, *ibid.*, 52, 1956, 65-71; E. D. Longenecker, J. Krauskopf, and M. E. Bitterman, Extinction following alternating and random reinforcement, this Journal, 65, 1952, 580-587.

[2] Grant, Hake, and Hornseth, *op. cit.*, 1-5; D. A. Grant and L. M. Schipper, The acquisition and extinction of conditioned eyelid responses as a function of the percentage of fixed-ratio random reinforcement, *J. exp. Psychol.*, 43, 1952, 313-320; Hake, Grant, and Hornseth, Resistance to extinction and the pattern of reinforcement: III. The effect of trial patterning in verbal "conditioning," *ibid.*, 41, 1951, 221-225; Longenecker, Krauskopf, and Bitterman, *op. cit.*, 580-587; D. C. Nicks, Prediction of sequential two-choice decisions from event runs, *J. exp. Psychol.*, 57, 1959, 105-114; L. B. Wyckoff and J. B. Sidowski, Probability discrimination in a motor task, *ibid.*, 50, 1955, 225-231.

if the totals are independently selected, also in the runs that *a* and *b* form. Event *a*, for example, begins to appear in longer runs than it did before, and these new run-lengths may be regarded as new 'run-stimuli.' We may, in fact, regard a shift in probabilities as if it were a more familiar transfer-problem, in which *S*s meet 'new' or 'old' stimuli and can carry over 'old' responses that may or may not be appropriate. To determine whether this type of approach is a fruitful one, we have asked three questions.

(1) Suppose we hold constant the change in over-all probability (always 50:50 shifting to 75:25), but vary the runs that *S*s meet and to which they learn a response in the 50:50 phase. Will this affect the speed with which *S*s learn to favor the 75% alternative?

(2) How do *S*s meet the change from 50:50 to 75:25? Does the response of predicting the 75% alternative show a smooth increase as we progress from trial to trial? Or are there several changes, so that to some runs this response shows a rapid increase and to others a much slower shift? If the rapid shift occurred with new run-stimuli and the slow shift with old run-stimuli, then it would not be so puzzling that *S*s shift more easily from, say, 15:85 to 70:30 than they do from 50:50 to 70:30.[3]

(3) Can we predict performance by individuals, rather than by groups, from one pattern of events to another?

The first two of these questions are considered in Section I; the third question is dealt with in Section II.

I. Transfer as a Function of Structure of Runs

S was asked to predict on each of 192 trials whether *E* would turn up a blue or a yellow card. For the first 96 trials, the proportion of blue (*b*) to yellow (*y*) was 50:50; for the remaining 96 trials, it was 75:25.

The 75:25 stimulus-series is the same for all groups. Within it, the most frequent runs are y_1, b_1, b_2, b_3, and b_4; the probability of a *b* occurring after each of these is close to 0.75. The frequency with which these runs occur and the proportion of times that they are each followed by *b* are given in Table I.

Where the groups differ is in the nature of the preceding 50:50 series. Group 1 (which we shall call the short-run or the 0.25 group), experiences no run of *b* or *y* that is longer than 4; the probability of a *b* occurring after a *b* is 0.25; of a *y* after a *y*, also 0.25. Of the runs that *S* will encounter most often in the 75:25 phase,

---

[3] W. K. Estes and J. H. Straughan, Analysis of a verbal conditioning situation in terms of statistical learning theory, *J. exp. Psychol.*, 47, 1954, 225-234; Allen Parducci, Alternate measures for the discrimination of shift in reinforcement ratio, this Journal, 70, 1957, 194-202.

the *S* in this group is pretty well limited to experience with y_1, b_1, and b_2. To the run y_1, *S* will have learned to predict most often a *b* — a response appropriate for the 75:25 phase. To the other runs, he will have learned to predict most often a *y* — inappropriate. This group we expect to be the slowest in shifting to a higher over-all percentage of predictions of *b*.

For Group 2 (the 0.50 or normal-run group), $p(b/b)$ and $p(y/y)$ during the 50:50 phase are 0.50. All of the runs in the later 75:25 phase are met in this 50:50 phase, and the response to them of predicting *b* approximately half the time will be neither grossly inappropriate nor particularly apt for the 75:25 phase. Group 3 (the 0.75 or long-run group) should have the greatest advantage. To all runs that include *b*, *S* will already have learned to predict *b* most of the time; of the runs to

TABLE I

Most Common Runs, Their Frequency and the Events that Follow Them in the Initial (50:50) and Second (75:25) Phases

Run	Probability that *b* follows a run				Frequency of run			
	50:50			75:25	50:50			75:25
	Group 1	Group 2	Group 3		Group 1	Group 2	Group 3	
y_1	.75	.50	.17	.78	36	24	12	18
b_1	.25	.50	.83	.78	36	24	12	18
b_2	.22	.50	.80	.79	9	12	10	14
b_3		.50	.75	.73	2	6	8	11
b_4		.67	.83	.75	1	3	6	8

which he has learned to predict *y* most of the time, only the stimulus y_1 will be occurring with any great frequency.

To build the stimulus-series for these groups, we determined the distribution of runs needed to give the closest approximation to the conditional probabilities that we wanted, and then placed these runs in a random sequence, alternating runs on *b* and *y*.

Subjects. The *Ss* were 108 U. S. Army enlisted men, ranging up to sergeant in rank. Time spent on the experimental task counted as a substitute for attendance at a session of the Troop Information Program (normally compulsory). Approximately half of the *Ss* were volunteers; the others were assigned to the task. *Ss* were run in groups of 10 to 15 men at a time.

Procedure. Ss were told that they were taking part in an experiment on "how people make judgments or predictions when they have limited information and limited time." Their job was to predict as accurately as they could whether *E* would turn up a blue or a yellow card. *E* advised them: "I cannot tell you how the stack is arranged, so at the beginning you will have to start off by guessing and see how your guesses work out."

Ss were given a booklet with pages numbered from 1 to 250. They wrote each prediction (B or Y) on a separate page, *E* calling before each trial the number of the page they should be on. After each prediction was written, *E* turned a card, held it up, and announced its color. The 192 trials (50:50 and 75:25) formed a continuous series, with no breaks in time once the series started.

RESULTS

(1) Frequency of predictions of blue (75%-event). The over-all speed of shift to more frequent prediction of blue in the 75:25 phase follows our expectations. Group 3 shows the most rapid increase in predictions of blue, then Group 2, and finally, Group 1. Table II shows these differences and the way they disappear as the 75:25 series continues.

We have included in Table II a measure of variability among trials and among *S*s. There has been recently much comment on variability among *S*s.[4] In the present circumstances, variability among trials is frequently

TABLE II

CHOICES OF BLUE (75% ALTERNATIVE) IN THE 75:25 PHASE AND IN THE LAST THIRD OF THE PRECEDING 50:50 PHASE

Phase and Trials	Mean choices of *b*			*SD* (trials)			*SD* (*S*s)		
	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
50:50 Phase:									
Trials 65–96	.53	.56	.48	.17	.16	.31	.09	.10	.07
75:25 Phase:									
Trials 1–32	.47	.56	.61	.16	.12	.21	.13	.14	.12
33–64	.63	.60	.63	.17	.14	.23	.11	.14	.10
65–96	.71	.71	.72	.13	.13	.15	.14	.18	.14

larger than among *S*s, an effect that underscores our concern with the specific stimuli on particular trials.

(2) Predictions of blue after runs of different length and pattern. Fig. 1 shows the main sources of predictions of *b*. Group 3 (the 0.75-group), for example, derives its advantage from carrying over an appropriate response to the stimuli b_1, b_2, b_3, and b_4, and is held back by carrying over an inappropriate response (predicting few *b*s) to the stimulus y_1. In contrast, Group 1 (the 0.25-group) is greatly handicapped by carrying over inappropriate responses to all runs but y_1. Group 2 (the 0.50-group) falls in between; it is somewhat better off than we expected, and is helped by the response developed in the 50:50 phase of over-predicting the occurrence of *b* after any run containing *b*. This group showed the same tendency toward over-predicting *y* after runs containing *y*, but these stimuli occur less often and are of far less moment in the 75:25 phase.

Fig. 1 suggests two interesting possibilities: (a) The same arrangement of responses to different stimuli appears to be carried over from the 50:50

[4] See Wyckoff and Sidowski, *op. cit.*, 225-231; Ward Edwards, Reward probability, amount, and information as determiners of sequential two-alternative decisions, *J. exp. Psychol.*, 52, 1956, 177-188.

phase to the first part of the 75:25 phase (*i.e.* the relative extent to which *Ss* predict *b* after the various runs seems to have considerable consistency). (b) The ease of shift appears to be related to the frequency with which a run has been encountered in the 50:50 phase and is now met with in the 75:25 phase. Group 1 (the 0.25-group) is a good example. This group learns to predict *b* more often after b_2, than after b_3 and b_4, while the predictions of *b* after b_1 lag behind. For the moment, we may note that of the stimulus-runs containing *b*, b_1 is the stimulus most frequently experienced in the 50:50 phase (where it was followed by *y* 75% of the time); of the

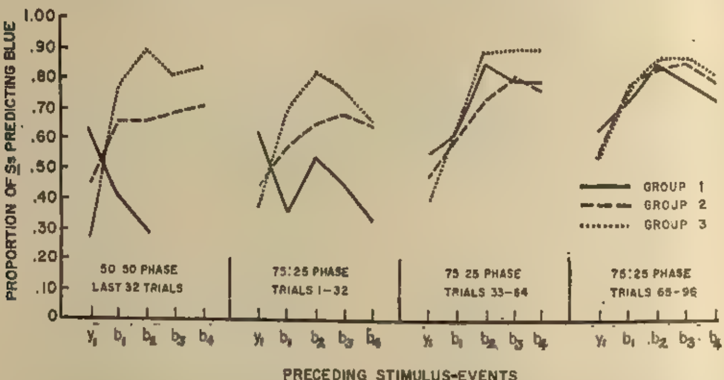


FIG. 1. RESPONSE TO VARIOUS LENGTHS OF RUN DURING INITIAL (50:50) AND SECOND (75:25) PHASE.

relatively new runs, b_2 is the one about which evidence of its now being followed most often by *b* accumulates most rapidly.

(3) Ease of shift with different length of run. As a check on the effect of runs of new lengths, we have used one of Parducci's situations.[5] His 40 *Ss* predicted whether an *a* or *b* would be called. For the first 120 trials, the proportion of *a* to *b* was 15:85. For the second 120 trials, it was 70:30. This extreme shift introduced several new runs and gave rise to an interesting phenomenon; namely, that the *Ss* shifted more easily from 15:85 than did the *Ss* whose initial experience was with 50:50. We ask, does the advantage of the 15:85 group come from its respose to the new runs?

Fig. 2 indicates that this was the case. It was in response to the runs

[5] Parducci, *op cit.*, 194-202. The data are from Parducci's Experiment I; the analysis is ours.

seldom or never experienced before that the 15:85 group had its advantage, particularly in the first two parts of the 70:30 phase.

As is usually the case in transfer problems, however, a simple specification of 'new' or 'old' conditions is not quite enough. We need to make first a reservation about the number of new runs. If only one new run appears, S apparently extrapolates from stimuli very close to it. For example, if the only new run is b_4, S will respond to it with an extrapolation from the responses he has learned to b_3 and b_2 (cf. Fig. 1 and, to anticipate, Fig. 3). In such a case, particularly rapid learning with the new run does not

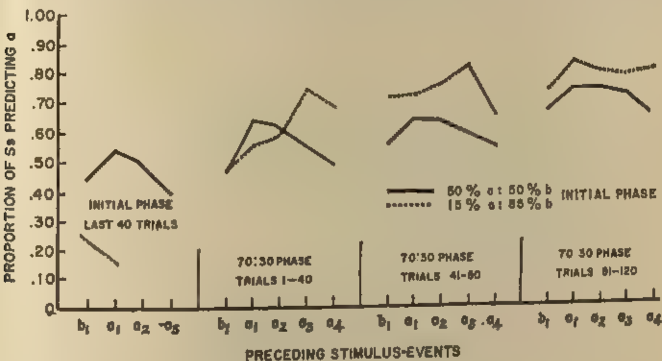


FIG. 2. DIFFERENCES IN RESPONSE TO NEW AND OLD STIMULI, FROM AN EXPERIMENT BY PARDUCCI.

occur. It is when S is faced with several new runs (*e.g.* b_2, b_3, and b_4 are all new) that he shows rapid learning and starts giving a response quite different from the one he had learned and continues to give to b_1. Apparently several new stimuli of this kind are too remote from b_1 to allow generalization from it.

The interesting point is that S, in such a situation, seems to look elsewhere for something from which to generalize, and he finds it in past runs of the same length as the ones he now faces but occurring on the opposite event. We base this suggestion on a comparison of the 15:85 group with an additional group provided by Parducci, a group starting off with 70:30 (no previous experience). For the runs the 15:85 group had experienced before, and to which it had learned an inappropriate response, the 15:85 group lagged behind the no-experience group. (During the first 40 trials of 70:30, the proportion of times the 70% side was predicted by the 15:85 and the no-experience group was respectively: after b_1, 0.46 as

against 0.61; after a_1, 0.56 as against 0.67.) For the stimulus a_2 (seldom experienced before), the 15:85 group started to draw ahead (0.58 against 0.50), and its lead became more pronounced with the stimuli a_3 and a_4 (0.74 against 0.41, 0.68 against 0.27).

Clearly, the 15:85 *Ss*, in their response to these new runs, are not behaving like *Ss* for whom 70:30 is the first experience. The most reasonable explanation is that they are generalizing from their previous experience with stimuli of the same length on the opposite event; *i.e.* they

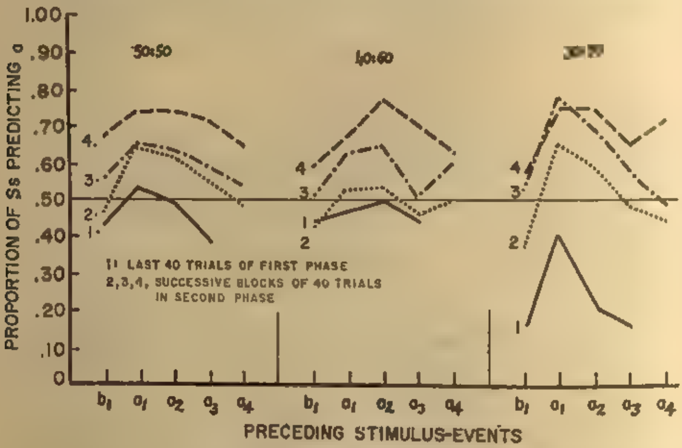


FIG. 3. RESPONSE TO VARIOUS STIMULI IN FIRST AND SECOND PHASES, FROM AN EXPERIMENT BY PARDUCCI.

predict another *a* after the news runs a_3 and a_4, just as they predicted *b* after b_3 and b_4 in the 15:85 group.

Apparently, then, what counts is not only the appearance of runs for which *S* has had little or no previous opportunity to establish an inappropriate response. There is also the question of *Ss* drawing on a response to runs that are perceived as somewhat like those previously experienced. This argument for stimulus-generalization is, of course, highly reminiscent of Sheffield's arguments, with the difference that we have defined the stimuli in a radically different manner.[6]

(4) Carrying over the same arrangement of responses. For those situations where much the same runs occur in the two phases, the *Ss* seem to carry over the same arrangement of responses to these stimuli, *i.e.* the

[6] V. F. Sheffield, Extinction as a function of partial reinforcement and distribution of practice, *J. exp. Psychol.*, 39, 1949, 511-525.

relative extent to which they predict an alternative after the various runs seems to remain the same. This we noted in Fig. 1. Fig. 3 provides additional evidence of this effect in three of Parducci's situations.[7] The groups shown in Fig. 3 all experienced the same 70:30 schedule in the second phase. They differed in their initial experience (50:50, 40:60, and 30:70), and each tended to carry over into the second phase the pattern of responses established to runs in the first phase.

II. The Prediction of Transfer for Individuals

Can performance be predicted from one stimulus-series to another, not by groups but by individuals? To answer this question we shall inspect the correlations between performance at one stage of learning and performance at another. This requires (1) a large sample of *Ss*, and (2) the use of more than one group of *Ss* to allow cross-validation. We used three groups of *Ss* (designated here as A, B, and C), all of which faced a shift from 50:50 stimulus-series to a series with unequal proportions. These groups differ somewhat as to procedure and are not drawn from the same population.

The *Ss* were asked to predict on each of 216 trials whether *E* would turn up a blue or a yellow card. For the first 120 trials the proportion of blue (b) to yellow (y) was 50:50; for the remaining 96 trials, it was 75:25. This held for both Groups A and B. In Group B, however, 14 additional cards, all blue, were added. Since the last 8 cards of the 216 trials also were blue, this made a total of 22 consecutive blue cards.

The stimulus-series were similar but not identical for Groups A and B. For Groups A and B the probability of b appearing after a b, that is, $p(b/b)$ was 0.47 in the 50:50 phase; $p(y/y)$ equaled 0.45. In the 75:25 phase, for Group A, $p(b/b)$ equaled 0.80 and $p(y/y)$ equaled 0.42. For Group B, in the 75:25 phase, $p(b/b)$ equaled 0.75 and $p(y/y)$ equaled 0.25. There were 67 *Ss* in Group A, 63 *Ss* in Group B. All of the *Ss* were U.S. Army enlisted men. Group A was tested in subgroups of about 20. Group B was tested in subgroups of 10 to 15. Group A was relieved from company duty and assigned to take the test. For Group B, time spent on the test substituted for attendance at a Troop Information Program session (normally compulsory). Approximately half of Group B were volunteers; the others were assigned.

As an additional check on results obtained with Groups A and B, we analyzed raw data provided by Parducci for a group of 38 college students.[8] The procedure was quite similar, *Ss* predicted on each trial whether the letter *a* or the letter *b* would occur: For the first 120 trials the proportion of *a* to *b* was 50:50; for the next 120 trials, 70:30.

---

[7] Parducci, *op. cit.*, 194-202. Parducci's Experiments I and II; our analysis.
[8] Parducci, *op. cit.*, Experiment I.

Results. Consider first the consistency of *Ss'* response to the preceding stimulus-runs in the two phases, for the three groups.

As a measure of *Ss'* response to runs, we shall take the proportion of times that they predict the stimulus that appeared on the preceding trial, $p(X/x)$.[9] This measure reflects essentially the extent to which *S* predicts that runs will continue; that is, that a run will be longer rather than shorter. If $p(X/x)$ were zero, *S* would be predicting an alternating series with runs of length 1.

Table III presents coefficients of correlations showing the consistency of $p(X/x)$ by *S* from the first to the second phase. For all three groups, these correlations are sizeable (0.47, 0.54, 0.53). The correlations in the

TABLE III

COEFFICIENTS OF CORRELATION BETWEEN REPEATED RESPONSES DURING INITIAL PHASE (50:50) AND DURING TWO PARTS OF SECOND PHASE (75:25) OR (70:30)

Response and previous event	Trials 1–40 (or 1–48)			Trials 41–96 (or 49–120)		
	Group A	Group B	Group C	Group A	Group B	Group C
$p(X/x)$*	.47	.54	.53	.26	.51	.34
$p(X/X$ correct)	.50	.77	.59	.34	.74	.54
$p(X/Y$ wrong)	.60	.75	.19	.38	.74	.17

* Upper case symbols refer to responses; lower case to stimuli.

next two runs are obtained by considering (a) the value of $p(X/x)$ when the *S* predicted *x* correctly, and (b) the value of $p(X/x)$ when the *S* made an incorrect prediction of *x*.

For Groups A and B, this consistency is maintained whether the previous prediction was correct or not. For Group C (Parducci's group), there is consistency when a stimulus had just been correctly predicted, but not when a stimulus had been incorrectly predicted. We do not know why this difference should occur.

How does the measure $p(X/x)$ predict the game score, *i.e.* the frequency with which the *S* chooses the more probable alternative in the 75:25 phase? We should expect that the *S* who predicts longer runs during the 50:50 phase will be, in general, better off when the series shifts to 75:25 or 70:30. He is, in fact, better off—the correlations between $p(X/x)$ in the 50:50 phase and the game-score in the first half of the second phase are 0.38, 0.32, and 0.28. Clearly, however, these correlations are not of a size with which one can be contented.

[9] Lower-case refer to stimuli; upper-case to responses.

Part of the difficulty, we suggest, lies in the lack of tight relationships between our measure of response to run-stimuli in the first phase, $p(X/x)$, and the performance that produces a game-score. We base this suggestion on finding higher correlations when the relation can be tightened. In Group B we added to the end of the series, 14 trials in which a blue card always turned up. Since blue also turned up on the last eight trials of the regular 75:25 series, this gave a run of 22 blues. In these 22 trials there are only two things that can happen to an *S*. He can predict *b* and be correct; or he can predict *y* and be wrong. To acquire a high game-score (high prediction of *b*), there are only two things to do: repeat a correct prediction of *b*, and switch over to *b* after an incorrect prediction of *y*. If we now measure the predictions of *b* during these 22 trials against the strength of these two responses in the last half of the 75:25 phase (before the run of 22 begins) the correlations are more satisfactory in size (0.49 and 0.44 respectively).

This increase is certainly only suggestive, being based on only one group, but it does encourage the belief that predictions by subject to the game-score can be satisfactorily made if we tighten experimentally the relation between the performance that produces the game-score and the response to specific stimuli during the preceding phase. As it is, the correlations lend weight to the argument that the *S*s tend to carry over, to runs that they have met before, the responses established in the first phase.[10]

Discussion

The results raise two questions: (1) about the definition of stimulus-runs in probability tasks; and (2) about the relation between our emphasis on runs and other hypotheses about response to change in the probabilities of events.

For probabilistic tasks in general, the results suggest that it is not enough to define stimulus-runs solely in terms of the absolute probability with which they occur. One 50:50 situation, for example, can be quite different from another as far as the structure of runs is concerned, and in fact is quite likely to be different since different experiments require

[10] We can also state that some other differences between individuals are not significant. The *S*s are not simply carrying over a preference for one alternative (correlations insignificant for all three groups), and the speed with which they make the shift is not consistently related to intelligence. Correlations with the five subtests of the Army Classification Battery for Group A are 0.12, 0.00, −0.01, −0.19, and −0.03; for Group B, 0.51, 0.46, 0.38, 0.39, and 0.44. We do not know why this discrepancy occurs, but the inconsistency alone inclines us against intelligence as a major variable.

a variety of ways of randomizing a stimulus-series. When the requirement, for instance, is that the 50:50 rule apply in each block of 10 trials, the run-structure will be different from the one in which the 50:50 rule applies over the total series. If *Ss* are responding to the structure of the run, such differences in stimulus-series make it difficult to compare results in terms of absolute probability alone.

On the positive side, knowledge of the structure and length of runs may give us a more direct approach to the problem of what determines the frequency with which an alternative, particularly a more probable alternative, is chosen. For example, in a $75b{:}25y$ situation, we find *Ss* predicting b around 75% of the time, and it is not clear how this matching is produced. Our results suggest that this could be a product of *Ss* heavily over-predicting b after b_1 and after b_2, but under-predicting it after, say, b_3 and particularly after y_1. The general task is shifted to asking where the *Ss*' predictions of length or structure of runs depart from the objective characteristics of runs he has just experienced. This is much the same question that has been asked about where discrepancies occur between subjective and objective probabilities with particular sets of betting odds.[11] This approach is also represented in the work of Lawlor and of Cohen and Hansel,[12] both of which provide considerable data on the structure of runs that *Ss* expect.

How does the emphasis on runs fit with other hypotheses about response to changes in probabilities of events? We do not propose that length of run is the only factor affecting response. More likely several processes are involved in such shifts. This extends Grant and Schipper's proposal that at least two parameters are involved: habit strength at the time of extinction and the discriminability of the extinction from the acquisition phase.[13] We suggest four processes:

(1) The discrimination that a change has occurred. It is probably here that the appearance of new stimuli, or the sudden and extreme inappropriateness of a previous response are most important.[14]

(2) The expectation that the change will be permanent. This is one of the conditions that Parducci stresses,[15] and it seems likely that this is where experience of previous shifts is most effective.[16]

---

[11] Ward Edwards, Probability-preferences in gambling, this JOURNAL, 66, 1953, 349-364.

[12] W. G. Lawlor, Subjective probability in sequential uncertainty situations. Unpublished Ph.D. thesis, University of Chicago, 1956; John Cohen and C. E. M. Hansel, *Risk and Gambling*, 1956.

[13] Grant and Schipper, *op. cit.*, 313-320.

[14] See Longenecker, Krauskopf, and Bitterman, *op. cit.*, 580-587.

[15] Parducci, *op. cit.*, 202.

[16] Hake, Grant, and Hornseth, *op. cit.*, 221-225.

(3) Learning what the new pattern or sequence is. It is here that most of the problems about discriminability arise. It has often been proposed that the discriminability of the new pattern or sequence is the critical variable but, as has also been pointed out, we are still lacking any independent measure of discriminability or similarity.[17] It has been proposed, however, that one condition affecting it is the presence of interference from the *S*s' interest in his own responses and his rewards.[18]

(4) Replacing the old response with something else. This, and the previous factors, may present no problem if *S* learns in the initial phase a form of response that is highly appropriate to the later phase. An *S*, for example, who learns to predict the stimulus that appeared on the previous trial, has no difficulty in shifting from, say, 100:0 to 0:100 and there is evidence that some experimental conditions promote this form of response.[19] Similiarly there is little problem if the new pattern is completely predictable, is easily discriminable, can be neatly coded (*e.g.* 100:0 or 0:100) and is expected to be permanent. It is when these conditions do not apply that we seem to be most sharply faced with the slow modification of old responses and the building up of new ones and with such variables as the length of the first acquisition series.[20] It is in this type of task that the problem of specifying the qualities of stimuli and responses becomes most acute.[21]

Summary

This study explores the value of defining stimuli in terms of length of run, for situations in which *S* faces a shift in the probability of an event (stimulus or reward). Shifts in probability are treated as transfer situations, in which *S* may meet *new or old* lengths of runs and can carry over old responses that may or may not be appropriate.

---

[17] Grant and Schipper, *op. cit.*, 313-320; D. W. Tyler, E. C. Wortz, and M. E. Bitterman, The effect of random and alternating partial reinforcement on resistance to extinction in the rat, this JOURNAL, 66, 1953, 57-65.

[18] J. S. Bruner, E. H. Galanter, and M. A. Wallach, The identification of recurrent regularity, this JOURNAL, in press; Goodnow and Pettigrew, Some sources of difficulty in solving simple problems, *J. exp. Psychol.*, 51, 1956, 385-392; A. J. Riopelle, Rewards, preferences and learning sets, *Psychol. Rep.*, 1, 1955, 167-173.

[19] Goodnow and Pettigrew, *op. cit.*, 1955, 381-389.

[20] Grant and Schipper, *op. cit.*, 313-320; D. J. Lewis and C. P. Duncan, The effect of partial reinforcement and length of acquisition series upon resistance to extinction of a motor and a verbal response, this JOURNAL, 69, 1956, 644-646.

[21] See Friedes, *op. cit.*, 385-392; Logan, Beier, and Kinkaid, *op. cit.*, 65-71; Grant and Schipper, *op. cit.*, 313-320; Tyler, Wortz, and Bitterman, *op. cit.*, 57-65.

SERIAL POSITION AND THE MEMORY-SPAN

By NANCY C. WAUGH, Massachusetts Institute of Technology

The memory-span has been measured by a method of constant stimuli, but traditionally the span is the average length of a series of stimulus-items recalled correctly in 50% of the trials, and both the number and the location of errors in incorrectly recalled series are beside the point.[1] An alternative way to define the span is to say that it is simply the number of consecutive items recalled correctly after one presentation. So defined, it can consist of a subseries of items within a longer series, measured from any arbitrary reference-point. If the reference-point is the first or the last item in a series, the span can be determined by a method of first errors, whereby it is based on the relative frequencies with which the first, the first two, and the first n items—or, conversely, the last, the last two, and the last n items—in the series are recalled correctly. The span is the mean number of items recalled correctly in an unbroken sequence; it is defined by the mean serial position of the first or the last item that is recalled incorrectly.

The span may thus be estimated from either of two cumulative distribution-functions: either the relative frequency with which an item and all of the items before it are recalled correctly, or the relative frequency with which it and all of the items after it are recalled correctly, as a function of this item's position in the series. These functions will be referred to here as the *initial* and the *terminal span*, respectively. Their values for the nth item from the beginning and from the end of a series must, of course, be equal to each other when the series consists of n items, and both must then be equal to the relative frequency with which the entire series is recalled correctly.

To each function, then, there corresponds a memory-span: the initial and the terminal spans. According to the method of first errors, the initial span is the mean number of items recalled correctly before the first failure in recall from the beginning of a series, while the terminal span is the mean number recalled correctly before the first failure in recall from the

* Received for publication October 2, 1958. The work reported in this paper was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology with the joint support of the U.S. Army, Navy, and Air Force.

[1] J. P. Guilford and K. M. Dallenbach, The determination of the memory span by the method of constant stimuli, this JOURNAL, 36, 1925, 621-628.

end of a series. These spans can also be determined by the method of constant stimuli if *S* is asked to recall initial or terminal sub-series of fixed lengths; the span will be the mean length of sub-series correctly recalled. Clearly, *S* need not recall an entire series for one or the other span to be determined by either method. Thus, the memory-span as it usually has been defined becomes a special case of the initial or terminal span, one which occurs when *S* has had to reproduce an entire series and when both spans cover the same set of items.

The initial and terminal span-functions may be said to represent the classical effects of primacy and recency, respectively. The span-function refers, it is true, to the recall of strings of items, whereas the effect of serial position has usually been measured by the recall of single items. If a correct response never occurred after the first failure in recall from either the beginning or end of a series, then the span-functions would coincide with the traditional curve of serial position. Insofar as this is not the case, the span-functions should lie farther below this curve for items farther away from the boundaries of a series.

Primacy and recency refer to the fact that items at the extremes of a series are easier to learn than those in the middle. The longer the series, the greater this discrepancy seems to be.[2] This finding suggests that the classical memory-span is limited primarily by the difficult middle items in longer series, as Foucault has implied,[3] and thus that the function ought to vary systematically with the length of the series in which they are measured, becoming steeper, at least up to a point, as the series becomes longer. This inference is consistent with the common observation that as a series of digits or letters becomes longer, the total number of items recalled correctly rises to a maximum and then falls. For example, Gates has reported that if a series exceeds the classical span by as few as four items, the number of items recalled correctly from this series is less than the span.[4] Similar results have been obtained by Pollack over a wide range of series-lengths.[5]

---

[2] J. J. Gibson and G. A. Raffel, Techniques for investigating retroactive and other inhibitory effects in immediate memory, *J. gen. Psychol.*, 15, 1936, 107-117; C. I. Hovland, Experimental studies in rote-learning theory. VII. Distribution of practice with varying lengths of list, *J. exp. Psychol.*, 27, 1940, 271-284; E. R. Robinson and M. A. Brown, Effect of serial position upon memorization, this JOURNAL, 37, 1926, 538-552.

[3] Marcel Foucault, Les inhibitions internes de fixation, *Année Psychol.*, 29, 1928, 92-112.

[4] A. I. Gates, The mnemonic span for visual and auditory digits, *J. exp. Psychol.*, 1, 1916, 393-403.

[5] Irwin Pollack, The assimilation of sequentially encoded information, this JOURNAL, 66, 1953, 421-435.

The following experiments deal with the relation between series-length and the span-functions.

Experiment I

The first experiment assesses the effect of series-length on the terminal span and compares the method of constant stimuli and the method of first errors for determining the initial and terminal functions.[6]

Method. The material to be learned consisted of eight digits (*O* through 7) arranged in quasi-random series in which no more than two successive digits were identical and in which blocks of 64 digits contained equal numbers of each digit. The items were presented successively, by means of a mechanical sequence-timer, in an electric digit-display device,[7] at a rate of one digit every 2 sec., with an exposure-duration of 1 sec. They appeared as dark numerals, 1-⅛ in. high, on a light background through a window that was 1-½ in. high and 1 in. wide. The display-device was on a table directly in front of *S*, who sat at a comfortable viewing distance from it.

The task was to reproduce the last *N* digits in successive series of various lengths. The set of digits to be reproduced will be referred to as the *recall-series,* while the total series presented will be called the *exposure-series.* The difference between the exposure-series and the recall-series—the extra digits that appeared before the recall-series—will be referred to as the *pre-recall series.* The exposure-series must, of course, always have been at least as long as the recall-series.

A series began when *S* pressed a switch and it was terminated automatically by the appearance of a black 9 on a red background. *S*, who was seated alone in a soundproof room, was required to reproduce the designated recall-series on a response-card, guessing if necessary, at the end of each series. *S* could write down the items in any order he chose and could spend as much time as he wanted in recalling them. He was never informed as to whether his responses were correct.

The experiment consisted of 10 blocks of 10 trials each. Within a given block, *S* was to recall a specified number of terminal digits from 1 through 10. Each recall-series was preceded by a pre-recall series of 0, 5, 10, 15, or 20 digits. Within a block of trials, the various exposure-series—of which there were two of each length—were ordered randomly. *S* always knew the length of the recall-series but never that of the exposure-series. All the *S*s recalled the same series in the same ascending order, beginning with recall-series of one digit and ending with recall-series of 10 digits. The experiment was spread out over three sessions.

The *S*s were three women and two men, aged 25 to 30 yr., all of whom were college graduates.

Results. To test statistically for the effect of length of the exposure-

[6] Pollack and his co-workers recently have reported various values of the terminal span for series of different lengths, and they have called this span the running memory-span when the length of the series is unknown to *S*. The present study is primarily an analysis of the frequency-distributions from which such spans are estimated. See Irwin Pollack, L. B. Johnson, and P. R. Knaff, Running memory-span, *J. exp. Psychol.,* 57, 1959, 137-146.

[7] The Burroughs Modular Lenticular Digit Display.

series, the data were treated in the following manner. (1) The number of items recalled before the first failure in recall from the end of a series was determined for each *S* with each length of pre-recall series. These data, averaged over all lengths of recall-series, are presented in Table I, where the maximal value possible in each cell is 5.5 digits. A Friedman two-way analysis of ranks for these data revealed that the differences between lengths of pre-recall series were significant ($X^2 = 12.28$, $0.02 > P > 0.01$), while the differences between *Ss* were not ($X^2 = 6.52$, $0.10 > P > 0.05$).[8] These conclusions were borne out by an analysis of variance ($F = 7.57$, $P < 0.05$ for the differences between lengths of pre-recall series; $F = 1,26$, $P > 0.05$ for the differences between *Ss*). (2) A multiple-range test was then per-

TABLE I

MEAN NUMBER OF ITEMS RECALLED BEFORE THE FIRST FAILURE IN RECALL FROM THE END OF A SERIES, AVERAGED OVER RECALL-SERIES OF 1–10 ITEMS

S	Number of items in pre-recall series					Mean
	0	5	10	15	20	
1	5.50	3.60	3.70	3.95	3.65	4.08
2	4.65	4.30	4.25	4.30	4.30	4.36
3	4.50	3.90	3.95	4.40	4.25	4.20
4	4.50	3.90	3.95	3.65	3.65	3.93
5	4.95	3.50	3.45	4.05	4.25	4.04
Mean	4.82	3.84	3.86	4.07	4.02	4.12

formed on the means of the columns in Table I.[9] According to this test, the mean of the first column (no items prior to the recall-series) differs significantly from the other four, which do not differ significantly among themselves. The individual data for the pre-recall series of 0 and of 5, 10, 15, and 20 items were therefore pooled in order to compute the span-functions.

Fig. 1 shows the terminal span obtained by the method of constant stimuli and the method of first errors, averaged over 5 *Ss* and over pre-recall series of 5-20 items. For the method of constant stimuli—dashed line—$P(Z_{i,1})$ denotes the relative frequency with which all *i*-items were recalled correctly in recall-series of length *i*; each point is based on 40 observations, and the span is 5.6 digits. For the method of first-errors—solid line—$P(Z_{i,1})$ denotes the relative frequency with which the last *i*-items were recalled in recall-series of $i + 1$, $i + 2$, . . . , and 10 items,

[8] See Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences*, 1956, 166-169. I am indebted to J. E. Keith Smith for suggesting this statistical procedure, as well as others used later on, and also for suggesting the term *method of first errors*.

[9] D. B. Duncan, Multiple range and multiple *F*-tests, *Biometrics*, 11, 1955, 1-42.

for i ranging from 1-9. Thus $P(Z_{9,1})$ is based on 40 observations, while subsequent points are based on sample-sizes in increasing steps of 40, and therefore $P(Z_{8,1})$ is based on 80 observations, $P(Z_{7,1})$ is based on 120, and $P(Z_{1,1})$ is based on 360. The span is 5.4 digits.

The vertical lines in Fig. 1 represent bounds of one *SD* for each point on the method of first errors, where the variance was estimated across recall-series of $i + 1$, $i + 2$, . . . , and 10 items. The *SD*s for i equal to 8 and 9 were not computed, since these points are based only on two and on one length of recall-series, respectively. Since none of the constant-stimulus points in Fig. 1 falls above or below one *SD* from each method-

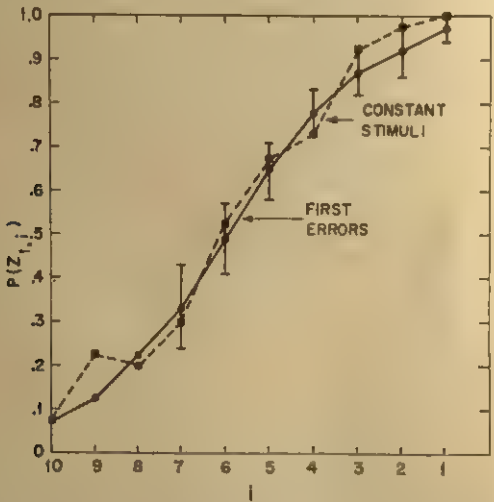


Fig. 1. Terminal Span-Functions Obtained by Two Methods
(The recall-series consisted of from 1-10 items and was preceded by 5, 10, 15, or 20 items.)

of-first-errors point, it may be argued that the method of constant stimuli and the method of first errors yield equivalent terminal spans when the recall-series is preceded by five or more items. This means, for example, that the last four digits in a series are recalled about as accurately when only these four must be recalled as when, say, these and the six that preceded them must be recalled; the recall of earlier digits does not seem to interfere with the recall of later ones when the pre-recall series consists of at least five items.

Fig. 2a shows the terminal span obtained by the method of constant stimuli and the method of first errors, averaged over 5 *S*s, when the

recall-series were not preceded by any other items. For the method of constant stimuli, each point is based on 10 observations. For the method of first errors, $P(Z_{9,1})$ is based on 10 observations, and subsequent points are based on sample-sizes in increasing steps of 10. The spans estimated from the constant method and the method of first errors are 8.5 and 7.4 digits, respectively. Note that the points on the curve obtained by method of first errors refer to items that were preceded by 0 to 9 other items in the recall-series, while the points on the curve obtained by the method of constant stimuli refer to items that were not preceded by any other items. There appears to exist no appropriate method for testing the statistical

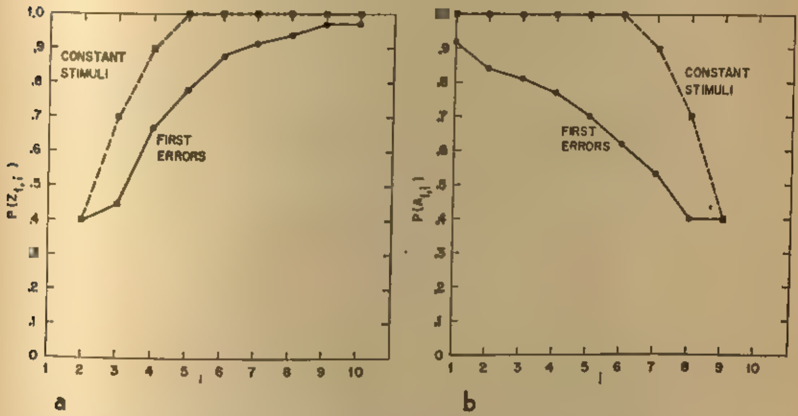


FIG. 2. SPAN-FUNCTIONS OBTAINED BY TWO METHODS
(The terminal and the initial span-functions obtained when the recall-series consisted of from 1-10 items and was not preceded by any other items.)

significance of the difference between the two functions, but their disparity suggests strongly that the last i-items in a series are more likely to be recalled correctly if item i is also the initial item in this series than if it is preceded by other items, and that primacy seems in this way to complement recency.

Fig. 2b shows the initial span-functions obtained with the same series as the terminal functions shown in Fig. 2a. Here $P(A_{1,i})$ denotes the relative frequency with which the first i-items were reproduced correctly when the recall-series coincided with the exposure-series. Every point is based on the same sample-size as its analogous point in Fig. 2a. The initial span obtained by the method of first errors is 6.4 digits. That obtained by the method of constant stimuli, 8.5 digits, is, of course, identical with the terminal span obtained by this method, and the initial

span-function obtained is the mirror-image of its terminal counterpart. The discrepancy between the two functions in Fig. 2b means that the first *i*-items in a series are likelier to be recalled correctly if item *i* is also the last item in this series than if it is followed by other items; and thus recency may be said to complement primacy.

Experiment II

The interaction between recency and primacy suggests that there exist two independent probabilities for the correct recall of an item in a series, which depend on its distance from the beginning and end of the series, respectively, and which interact noticeably when the series is relatively short. The next experiment was performed to determine how the span-functions for a relatively short series might be generated from the functions for a relatively long series on the basis of these assumptions.

Method. The apparatus and the general experimental procedure were the same as in Experiment I, except that the rate of presentation of the digits was increased to 1 per sec., with an exposure-duration of about 0.9 sec. The recall-series always consisted of 12 digits, while the exposure-series consisted of either 12 or 18. The tasks were to reproduce: (1) 10 recall-series that were the first 12 items in an exposure-series of 18 (Condition α); (2) 10 recall-series that were the last 12 items in an exposure-series of 18 (Condition ω); and (3) 10 exposure- and recall-series of 12 items only (Condition αω). The first two tasks were ordered randomly within the same blocks of trials; *S* was not informed as to which set of items he was to recall until he turned over the appropriate response-card at the end of each series. Interspersed with these series were 5 in which the recall-series was the first 6 digits and 5 in which the recall-series was the last 6 digits in the exposure-series of 18, which were included in order that items in each position in these series be requested equally often. *S* always was informed as to the length of the 12-digit series.

The *Ss* were seven women and three men, aged 23 to 32 yr., five of whom had served in Experiment I and five of whom had not. These groups will be designated as Group S (experimentally sophisticated) and Group N (experimentally naïve). Two blocks of each kind of series were presented in balanced order to each *S* over the course of two experimental sessions. The order was reversed for Groups S and N.

Results. The initial and terminal functions obtained with the 18-digit lists for Groups N and S are presented in Figs. 3a and 3b. These functions are based on both the 12-digit and the 6-digit recall-series. The terminal spans for Groups N and S, respectively, are 5.09 and 5.15 digits; the initial spans are 1.29 and 1.99 digits. Neither the difference between the terminal nor between the initial span-functions achieves statistical significance at the 5% level, according to the Kolmogorov-Smirnov two-sample

test.[10] Group S's longer initial span does, however, suggest that practice enhances primacy more than recency, a fact that has been noted by Lepley.[11]

The span-functions obtained with the 12-digit lists are represented by the dashed lines in Figs. 3c and 3d. The terminal spans for Groups N and

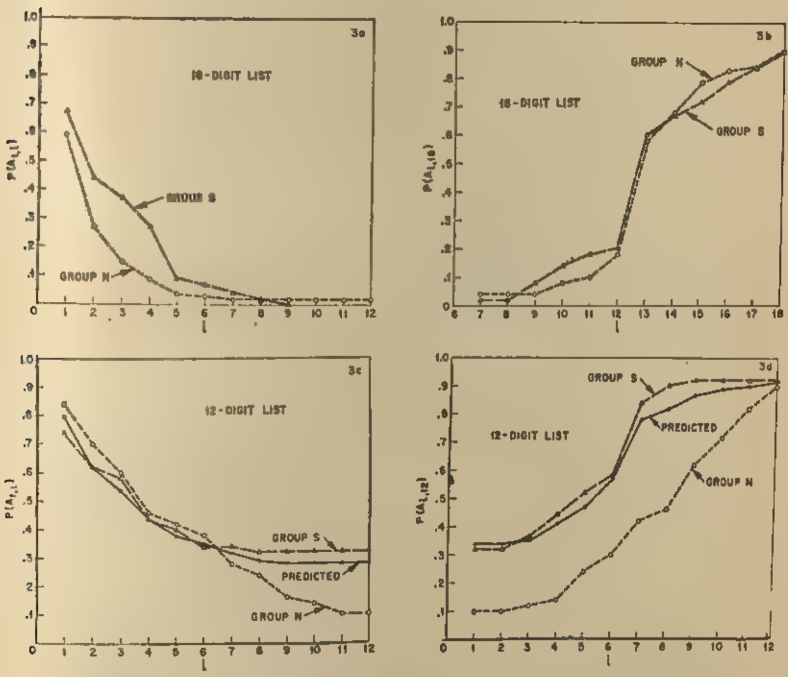


FIG. 3. SPAN-FUNCTIONS OBTAINED FOR TWO LENGTHS OF LIST
(The initial (a) and terminal (b) functions obtained with the 18-item list, and the initial (c) and terminal (d) functions obtained with the 12-item list.)

S respectively are 4.9 and 8.0 digits; the initial spans are 4.4 and 5.1 digits. A comparison of Fig. 3d with Fig. 3b reveals that Group N's terminal span-function for the 12-digit series lies below its terminal-span function for the 18-digit series. This anomolous fact may have resulted from the order in which the items were recalled: the *Ss* in Group N

[10] L. A. Goodman, Kolmogorov-Smirnov tests for psychological research, *Psychol. Bull.*, 51, 1954, 160-168; N. Smirnov, Tables for estimating the goodness of fit of empirical distributions, *Ann. Math. Stat.*, 19, 1948, 279-281.

[11] W. M. Lepley, Serial reactions considered as conditioned reactions, *Psychol. Monog.*, 46, 1934 (No. 205), 1-56.

tended to report that they recalled the items of the 12-digit exposure-series in order from first to last, while the *S*s in Group S tended to report that they wrote down late items before early ones.[12] A similar difference between the terminal functions for the longer series is not apparent, perhaps because the order in which the *S*s in Group S wrote down the recall-series then depended on its location.

Two hypotheses were formulated to describe the span-functions obtained with the 12-item list (Condition $\alpha\omega$) on the basis of the data obtained with the 18-item lists (Conditions α and ω). It was assumed in both cases that an item is recalled correctly under Condition α only by virtue of its proximity to the beginning of a series and is recalled correctly under Condition ω only by virtue of its proximity to the end of the series.

The first hypothesis states that the probabilities under Conditions α and ω of an item's correct recall, given that the items *before* it were recalled correctly, combine independently with each other under Condition $\alpha\omega$. The second hypothesis states that the probabilities under Conditions α and ω of an item's correct recall, given that the items *after* it were recalled correctly, also combine independently under Condition $\alpha\omega$. Neither of these hypotheses implies the other. Insofar as they both describe the present data, they suggest that an item in an unbroken string of correctly recalled items from either end of a series acts as a prompt for the next item in the string, and that the occurrence of the prompt and of the item prompted may depend on their distances from either or both extremes of the series.

These hypotheses may be stated in the following way:[13]

Let a_i, where $i = 1,2, \ldots n$, equal 1 when the ith item from the beginning of a series is recalled correctly and 0 when this item is not recalled correctly.

Let A_i denote that $a_i = 1$, and let $A_{i,j} (i \leq j)$ denote that the product $\Pi_{h=i}{}^{j} a_h = 1$. Thus $A_{1,j}$ signifies that all of the first j items in a series were correctly recalled, and $A_{i,n}$ signifies that all of the last $n - i + 1$ items were correctly recalled. Thus:

$$P(A_{1,i}) = P(a_1 = a_2 = \cdots = a_i = 1)$$
$$= P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_{1,2}) \cdots P(A_i \mid A_{1,i-1}) \quad [1]$$

and

$$P(A_{i,n}) = P(a_n = a_{n-1} = \cdots a_i = 1)$$
$$= P(A_n)P(A_{n-1} \mid A_n)P(A_{n-2} \mid A_{n-1,n}) \cdots P(A_i \mid A_{i+1,n}). \quad [2]$$

---

[12] For discussions of the relation between order and accuracy of recall, see J. Brown, The nature of set-to-learn and of intra-material interference in immediate memory, *Quart. J. exp. Psychol.*, 6, 1954, 141-148; James Deese and R. A. Kaufman, Serial effects in recall of unorganized and sequentially organized verbal material, *J. exp. Psychol.*, 54, 1957, 180-187; H. Kay and E. C. Poulton, Anticipation in memorizing, *Brit. J. Psychol.*, 42, 1951, 34-41.

[13] I am indebted to George A. Miller for this notation.

Let P' represent a theoretical probability. P_α and P_ω are based on the experimental data for long (18-item) series and $P_{\alpha\omega}$ is based on the data for short (12-item) series. The problem is to derive expressions for $P'_{\alpha\omega}$ based on the observed values of P_α and P_ω and then to test the result against the observed values of $P_{\alpha\omega}$.

If the value of $P_{\alpha\omega}$ for an item in a given position in a 12-item series represents a combination of the values of P_α and P_ω for the items in corresponding positions from the beginning and from the end respectively of an 18-item series, then $P_\alpha(A_1)$ should combine with $P_\omega(A_7)$ to determine $P_{\alpha\omega}(A_1)$ and $P_\alpha(A_{12})$ should combine with $P_\alpha(A_{18})$ to determine $P_{\alpha\omega}(A_{12})$. The conditional probabilities observed for items in other corresponding positions under Conditions $_\alpha$ and $_\omega$ likewise should combine, according to the present hypotheses. The first hypothesis of the combination of independent tendencies can thereby be interpreted to mean that, for the initial span-function,

$$1 - P_{\alpha\omega}'(A_i \mid A_{1,i-1}) = [1 - P_\alpha(A_i \mid A_{1,i-1})][1 - P_\omega(A_{i+6} \mid A_{7,i+6})]. \qquad [3]$$

The second hypothesis can be interpreted to mean that, for the terminal span-function

$$1 - P_{\alpha\omega}'(A_i \mid A_{i+1,12}) = [1 - P_\alpha(A_i \mid A_{i+1,12})][1 - P_\omega(A_{i+6} \mid A_{i+7,18})]. \qquad [4]$$

In the present case, $P_\omega(A_{i+6}|A_{7,i+6})$ and $P_\alpha(A_i|A_{i+1,12})$ were estimated by $P_\omega(A_{i+6}|A_{i+6})$ and $P_\alpha(A_i|A_{i+1})$, respectively, since the sample-sizes on which the former values are based become so small as i decreases or increases respectively that these probabilities eventually assume distorted values of 1 or 0. The Markov-like assumption behind these approximations can be justified by the fact that

$$P_\omega(A_{18})P_\omega(A_{17} \mid A_{18})P_\omega(A_{16} \mid A_{17}) \cdots P_\omega(A_i \mid A_{i+1})$$

and

$$P_\alpha(A_1)P_\alpha(A_2 \mid A_1)P_\alpha(A_3 \mid A_2) \cdots P_\alpha(A_i \mid A_{i-1})$$

were found to provide excellent approximations to $P_\omega(A_{i,18})$ and $P_\alpha(_{1,i})$, respectively.

In Figs. 3c and 3d, the predicted initial- and terminal-functions are represented by the solid lines. They were derived from the pooled data obtained from both groups with the 18-digit exposure-series. The Kolmogorov-Smirnov one-sample test was used to test the agreement between the observed and the predicted functions, since no exact method was available for this purpose. The predicted functions thus were treated as theoretical distributions, a procedure that would tend to overestimate the significance of the observed differences. Both the observed initial and terminal span-functions obtained for Group S are consistent with the hypotheses, as is the initial span-function obtained for Group N: in no case did the differences between observed and predicted functions attain statistical significance at the 5% level. The terminal span-function obtained for Group N, on the other hand, differs significantly from the predicted function. A possible reason for the low values of this observed function, the order in which the items were recalled, already has been mentioned.

DISCUSSION

The results of this study suggest that the memory span is determined by serial position effects. When a series is relatively short, recency and primacy complement one another, and a relatively large number of consecutive items are recalled from either end of a series. When a series is relatively long, recency and primacy cannot interact, and the terminal and initial spans are relatively short. The terminal span seems in this case to be independent of list length, at least over the range examined in Experiment I. Primacy and recency were denoted in Experiment II by the conditional probabilities for an item's correct recall.

It has been found that, when S is to recall an unspecified number of terminal items in a series, his terminal span is longer when he knows how long this series will be than when he does not. Apparently in this case he chooses an arbitrary origin for a recall series at some intermediate point in the exposure series, and primacy and recency thereby interact to determine the terminal span function. When S does not know how long an exposure series is going to be, on the other hand, and knows only that he is to recall a certain number of terminal items, as in Experiment I, he perhaps defines one recall series after the other and attempts to retain the most recent one, which may sometimes coincide with the one defined by E. The extent to which this strategy might differ from the strategy that S would pursue in attempting to retain an entire list merits investigation.

The hypotheses described in Experiment II would be consistent with a two-factor explanation of the effect of serial position. Such an explanation might, for instance, have recency result from the shorter delay between the exposure and the recall of the later items in a series, and primacy from the more intensive rehearsal of the earlier items. A more traditional explanation of serial position effects has been in terms of intraserial interference. Such interference is said to be exerted by the early items in a series on the later ones, and vice versa; according to Foucault, it increases progressively with distance from the beginning and end of a series. It

[illegible] Brown and Knoll, *op. cit.*, [illegible]

[illegible] *decay theory of immediate memory*, [illegible] *J. exp. Psychol.*, 10, 1958, 12[illegible]. R. Conrad and [illegible] A. Hille, The decay theory of immediate memory and paced recall, *Canad. J. Psychol.*, [illegible]

[illegible] The determinants of the effect of primacy [illegible] G. D. [illegible] and [illegible] primacy as a factor in [illegible] [illegible] and P. A. [illegible] VI. Intraserial interference, *J. exp. Psychol.*, [illegible]

Foucault, *op. cit.*, 92 [illegible]

that were the case, however, the frequency with which the first and the last items are recalled correctly ought to decrease with increasing series length, until finally neither these items nor any of the intervening items are recalled correctly at all. It is difficult, therefore, to see how P_a and P_m might be related to Foucault's proactive and retroactive inhibition. It is also difficult to relate the present hypotheses to theories that account for serial position effects just in terms of remote forward associations,[19] unless such theories can be made to state that the probability of recalling an item in a series is determined independently by its distance from each extreme of the series.

SUMMARY

The initial and terminal memory spans were defined as the mean number of consecutive items recalled correctly before the first failure in recall from the beginning and end of a series, respectively. In one experiment, the terminal span did not vary with the total length of the series presented when this length was unknown to S and when the sub-series to be recalled was preceded by five or more items. A method of constant stimuli and a method of first errors yielded similar terminal spans under this condition. When this sub-series was not preceded by any items, the values of the initial and terminal spans determined by the method of first errors tended to fall below the values of these functions determined by the method of constant stimuli. In a second experiment, the span functions obtained with a 12-item series were predicted by an hypothesis that asserts the independent combination of the conditional probabilities observed for the recall of item i from the beginning and item $13-i$ from the end of an 18-item series.

[19] B. R. Bugelski, A remote association explanation of the [illegible] learning [illegible] syllables in a serial list, [illegible] C. L. Hull, The conflicting psychologies of learning [illegible] 42, 1935, 491-516, [illegible]

THE TOLERANCE FOR PAIN AND FOR SENSORY DEPRIVATION

By ASENATH PETRIE, WALTER COLLINS, and PHILIP SOLOMON,
Harvard Medical School

Varied reactions to pain have long been noted by investigators. There is a tendency to assume that these variations are due to differences in the control of the experience, but the work discussed in this paper suggests that the differences are inherent in the experience of pain itself.

Just as there are great differences in visual sensitivity between blindness and the acute normal vision of youth, so sensitivity for pain can vary in degree. A man with incurable cancer can be relieved from pain by pre-frontal lobotomy, but the cause of the pain is still there, though he has ceased to suffer from it. Apart from the relief from pain that lobotomy brings, there are individual differences in algesic sensitivity. Some people sense pain acutely, others dimly, and still others scarcely feel it at all.

Sensory pain characteristically accompanies an excess of stimulation, whereas the stress of sensory deprivation (lack of stimulation) and monotony (lack of change in stimulation) are associated with a dearth of stimulation. Sensory deprivation is also borne differently by different persons; there are those for whom such deprivation creates great stress and also those for whom the stress is but minor.

The hypotheses of the present study are these. (1) Those aspects of personality that are changed by a pre-frontal lobotomy differentiate persons with high and low tolerance for pain. (2) Differences in tolerance for pain are paralleled by differences in perception. (3) Differences in tolerance for sensory deprivation are paralleled by differencs in perception that are the reverse of those associated with tolerance for pain.

METHOD AND PROCEDURE

Subjects. Altogether, 78 *Ss* were used in this study. They were divided as follows: 42 were patients undergoing different degrees and kinds of pain as a result of surgery or bronchoscopy; 19 were undergoing experimental pain; and 17 were undergoing sensory deprivation. The *Ss* undergoing experimental pain and deprivation were paid by the hour; the others were not paid.

---

* Received July 1, 1959, and accepted for prior publication December 29, 1959. This work was made possible by the active coöperation of the Boston Sanatorium and Beth Israel Hospital, Boston, with the collaboration of the Boston City Hospital. It was supported in part by the National Institute of Mental Health (M2641), the Lasker Foundation, and the Office of Naval Research.

Methods of assessing tolerance for pain and for deprivation. The patients in this study experiencing clinical pain were grouped according to their tolerance for pain (caused by comparable trauma) by three judges independently—a physician, a surgeon, and a nurse—who, in coming to their decisions, took into consideration (1) the patients' demands for analgesic drugs; (2) their sleeplessness as said to be caused by the pain; (3) the physical signs; and (4) the patients' statements of their experiences. The patients were interviewed by the physician on two occasions while they were in pain.

The *Ss* in the experimental group were subjected to pain caused by heat. Their thresholds for pain had been determined by Dr. Ulric Neisser,[1] with an adaptation of the Hardy-Wolff-Goodell dolorimeter, an instrument that concentrates radiant heat upon the skin.[2] The measurements are of temperature; the higher temperatures arise from longer exposure to the stiumulus. The temperature as which *S* first sensed pricking pain and the temperature at which he could no longer endure the pain were determined. For this second measurement, *S* was instructed to endure the pain as long as he possibly could. The difference between these two thresholds was used as a measure of *S*'s algesic tolerance—a new measure which we may call *tolerance for pain.* The group of experimental *Ss* was divided into thirds with low, medium and high tolerance for pain. The scores measuring tolerance for pain were about normally distributed in the population studied.

Tolerance for sensory deprivation and monotony was measured in 17 *Ss* who had volunteered to remain in a tank-type respirator under conditions inducing some sensory deprivation.[3] Our *Ss* took part in two investigations by Wexler, Mendelsohn, Leiderman, and Solomon.[4] The length of time that *S* was willing to remain in the respirator was used by us as the measure of his tolerance.

Tolerance for Pain in Relation to Personality

Broadly speaking, a person after pre-frontal lobotomy is more like a psychopath and less like a depressive. He is not a psychopath, but all his personality scores are changed in this direction. He might be said to be more 'extraverted.' Evidence for this statement comes from our studies of 100-odd patients with excisions of different extents in the pre-frontal regions of the brain.[5]

---

[1] Ulric Neisser, Temperature thresholds for cutaneous pain, *J. appl. Physiol.*, 14, 1959, 368-372.

[2] J. D. Hardy, H. G. Wolff, and Helen Goodell, *Pain Sensations and Reactions*, 1952, 67.

[3] The effects of sensory deprivation were first described by W. H. Bexton, Woodburn Heron, and T. H. Scott, Effects of decreased variation in the sensory environment, *Canad. J. Psychol.*, 8, 1954, 70-76. Subsequent experiments induced sensory deprivation by immersion in water; see J. C. Lilly, Mental effects of reduction of ordinary levels of physical stimuli on intact healthy persons, *Psychiat. Res. Reports*, 5, 1956, 1-28.

[4] J. H. Mendelson and J. M. Foley, A mental abnormality in poliomyelitis patients treated in a tank-type respirator, *Trans. Amer. neurol. Assoc.*, 81, 1956, 134-138; Donald Wexler, Jack Mendelson, P. H. Leiderman, and Philip Solomon, Sensory deprivation: a technique for studying psychiatric aspects of stress, *A.M.A. Arch. neurol. Psychiat.*, 79, 1958, 225-233.

[5] Asenath Petrie, *Personality and the Frontal Lobes*, 1952, 96-99; also, Effects of

In the present study of personality and its relations to tolerance for pain, we utilized the Maudsley Personality Inventory, the scores of which distinguish the psychopath from the depressive.[6] We used this inventory with three separate groups of patients at two Boston hospitals—the Boston Sanatorium and the Beth Israel Hospital. The patients had been subjected to (1) major chest surgery; (2) minor surgery, or (3) bronchoscopy. In addition, we used it (4) with volunteers subjected to experimental thermal pain and (5) on the first group of 10 vounteers who had been subjected

TABLE I

E-Scores on the Maudsley Personality Inventory for Five Classes of *Ss* further Divided into Sub-Groups on the Basis of Tolerance for Pain and Tolerance for Sensory Deprivation

Class of Stress	*N*	Most tolerant sub-group	Least tolerant sub-group	Direction of difference
Pain from:				
Major chest surgery	16	24.9*	21.3	+
Minor surgery	9	25.9	18.5	+
Bronchoscopy	17	27.4	26.5	+
Experimental pain	13	28.67	23.57	+
Sensory deprivation	10	28.50	31.80	−

* Higher E-scores characterize the patient after lobotomy and the more psychopathic patient. The difference between sub-groups is significant at the 5%-level.

to sensory deprivation. The *Ss* represented the extremes of tolerance and intolerance for pain; those moderately tolerant of pain were omitted from consideration.

In all five groups, our predictions were borne out. The scores are summarized in Table I. Results resembling those of a patient after lobotomy (the more 'psychopathic' scores) were found in the best tolerators, while those who suffered most had scores resembling a patient before lobotomy (the least 'psychopathic' scores). These relationships are reversed for the tolerance of sensory deprivation.

Perceptual Satiation Associated with Tolerance for Pain and Deprivation

With the experimental *Ss* and with the slightly ill patients, we were able to explore some of the perceptual characteristics of the *Ss* that are associated with the personality-type that tolerates pain. The important perceptual phenomenon in our study is the tendency, identified by Köhler and Wallach, for the intensity of a perception to be reduced in some persons

chlorpromazine and of brain lesions on personality, in H. D. Pennes (ed.), *Psychopharmacology*, 1958, 99-115.

[6] H. J. Eysenck, *The Dynamics of Anxiety and Hysteria*, 1957, 31.

after they have been stimulated for some time.[7] These investigators used the term *satiation* to describe this tendency and found individual differences in respect of it. It has been indicated that the susceptibility to satiation, as measured in one modality, like vision, is correlated with susceptibility to satiation in other modalities, like touch.[8]

The reduction in the subjective magnitude of a kinesthetic experience has, moreover, been shown to be related to the dimension of personality called *Introversion-Extraversion.* The more extraverted the personality, the greater the reduction.[9] The tendency toward reduced apparent size is also increased by certain types of brain injury.[10] Presumably the site of this phenomenon is central, not peripheral. The effect of brain-lesions is, however, apparently selective. Earlier work by Petrie shows that operations outside of the pre-frontal region which have no effect on sensitivity to pain, also have no effect on the tendency toward reduced size, nor any on some other measures related to extraversion.[11]

In this study we have measured satiability of kinesthetic size, adapting the apparatus first described by Köhler and Dinnerstein and subsequently used by other investigators.[12] *S* is first occupied for 45 min. answering questions, *and his hands are not used during this time.* We have introduced this resting period and consider it essential, as it permits the effect of whatever *S* handled prior to the testing session to wear off. Then *S* is blindfolded and feels with the thumb and forefinger of his right hand the width (38.1 mm.) of a test-object, a standard block of smooth, unpainted wood. Next, with the thumb and forefinger of his other hand, he feels a long tapered bar of similar unpainted wood and determines the place on the bar where it seems just as wide as the test-block. Movement back and forth along the bar is permitted, and the position of subjective equality for perceived width is thus fixed. The measurement is made four times in succession.

---

[7] Wolfgang Köhler and Hans Wallach, Figural after-effects, *Proc. Amer. phil. Soc.*, 88, 1944, 269-357. It should be made clear that certain of the after-effects reported by Köhler and Wallach are displacements that result in an enlargement of the apparent size of a figure. We are unable fully to confirm the results of Köhler and Wallach when *Ss* are grouped as Reducers and Non-reducers, in that the average effect on Reducers of stimulation with a large and a small block is to make them reduce while on the Non-reducers it is to make them enlarge. Details are being presented in a later paper. *Ss* of the two types should be separated in studying these two phenomena.

[8] Michael Wertheimer, Figural after-effect as a measure of metabolic efficiency, *J. Pers.*, 24, 1955, 56-73; Michael Wertheimer and Nancy Wertheimer, A metabolic interpretation of individual differences in figural after-effects, *Psychol. Rev.*, 61, 1954, 279-280.

[9] Eysenck, *op. cit.*, 156-159.

[10] G. S. Klein and David Krech, Cortical conductivity in the brain injured, *J. Pers.*, 21, 1952, 118-148.

[11] Petrie, Effects on personality of excisions in different regions of the brain with special reference to the relief of pain, *Proc. 14th int. Cong. Psychol.*, 1954, 167; Asenath Petrie, Walter Collins, and Philip Solomon, Pain sensitivity, sensory deprivation, and susceptibility to satiation, *Science*, 128, 1958, 1431-1433.

[12] Wolfgang Köhler and Dorothy Dinnerstein, Figural after-effects in kinesthesis, *Miscellanea Psychologica A. E. Michotte*, 1947, 196-220; Klein and Krech, *op. cit.*, 118-148; Eysenck, *op. cit.*, 155-156; Petrie, *Psychopharmacology*, 1958, 99-115.

S is then given a wider test-block (63.5 mm.) which he rubs with his right-hand finger and thumb at a constant rate for 30 sec. The purpose of this rubbing of a wider block is to induce satiation if it is to occur. After the rubbing, he again equates the original test-block to the perceived equivalent width on the tapered bar, determining four equivalences.

Next the time of satiation-rubbing is increased, first to 60 sec., then to 90, then to 120, and four measurements of the subjective size of the test-object are made after each period of satiation. Thereafter, instead of rubbing to induce satiation, an empty interval is allowed to elapse, with measurements of the test-object after 15 min.

The second group of deprived *Ss* were asked in addition to measure the larger block. They were given only one period of stimulation with the larger block and the whole process took 2 min.

We may call the persons who tend to reduce the size of the block subjectively after stimulation *Reducers,* contrasting them with the *Non-reducers* who accept the environment in its full intensity.

Our results support the hypothesis that the greatest tolerance for pain is shown by Reducers. We think that this tolerance for pain is partially due to the *Ss'* tendency to reduce the effectiveness of stimulation. Thus a brief wave of intense pain could cause later pain to appear less severe. Indeed, our own experiments on satiability indicate that this diminution is a cumulative process which persists much longer than the stimulus that causes the reduction.

Experimental pain was induced in 19 *Ss* by applying heat to the skin. Reducers were found to tolerate this pain best; Non-reducers poorly. The differences in satiability in the kinesthetic task between those who could and could not tolerate pain reached the 5% level of confidence after only 90 sec. of kinesthetic stimulation. An extreme Reducer diminished the apparent size of the wooden block by 30% after a further 90 sec. and this effect was still partially operative a quarter of an hour after all stimulation ceased (see Table II and Fig. 1). A pilot study of satiability in surgical patients who could sit up in a wheel chair also shows that tolerance for clinical pain is greatest in Reducers.

If we are right in explaining the tolerance of Reducers for pain as being partially due to a reduction in the intensity of stimulation, then this tendency to reduce should be a handicap in a situation where the environment starves the individual of sensory experience instead of bombarding him with it, as is the case for pain. Such sparsity of stimulation occurs in the experimental circumstances arranged to produce sensory deprivation or sensory monotony. Thus, being a Reducer should make for intolerance of sensory deprivation, because it would render the already limited stimulation even less effective.

Our findings on two different groups of 17 volunteers, all subjected to sensory deprivation, indicate that this relationship holds.[13] We find that it is Non-reducers who can better tolerate this starvation of stimulation as measured by the number of hours they are willing to remain in a tank-type respirator, while it is the Reducers who cannot for long accept this kind of stress. This difference is exactly the reverse of the behavior of these two types under the stress of pain. Results for the first group of 9 *Ss* are given in Fig. 2 and Table II. In the second group of volunteers, the difference in satiability between the good and poor tolerators of deprivation

TABLE II

REDUCTION IN APPARENT SIZE OF BLOCK FOR FIVE GROUPS OF *Ss* DIFFERING IN TOLERANCE OF PAIN AND OF DEPRIVATION (in millimeters)

Time of test	Pain: Least tolerant ($N=7$) Mean	SE	Pain: Moderately tolerant ($N=6$) Mean	SE	Pain: Most tolerant ($N=6$) Mean	SE	Deprivation: Least tolerant ($N=4$) Mean	SE	Deprivation: Most tolerant ($N=5$) Mean	SE
After 30 sec. stimulation	0.81	1.07	0.42	1.68	2.46	1.47	1.26	0.95	−0.03	0.86
After 90 sec. stimulation	−0.42[a]	0.77	1.92	1.69	3.50[a]	1.64	1.74	0.75	−0.12	0.83
After 180 sec. stimulation	−0.09[ac]	0.56	2.97	1.81	4.95[a]	1.71	4.77[bc]	1.11	1.86[b]	0.70
After 300 sec. stimulation	0.96[ac]	0.87	3.36	1.82	5.28[a]	1.96	4.68[bc]	1.23	2.04[b]	0.63
After 15-min. rest	−0.42[ac]	0.46	0.72	1.24	4.02[a]	1.35	2.55[c]	0.69	1.83	1.19

Differences are significant between: [a] Least and most tolerant of pain; [b] Least and most tolerant of deprivation; [c] Least tolerant of pain and deprivation.

is significant beyond the 2% level of confidence. Indeed, if the two deprivation groups are combined, the difference in satiation after 120 sec. of stimulation is significant beyond the 1% level.

Contrasting attitudes concerning their pain were, moreover, expressed by those *Ss* who were most and least tolerant of the stress of sensory deprivation. These two groups differed in their estimations of their tolerances for pain, in their estimations of their parents' attitudes towards the expression of suffering with pain, and in their evaluation of the pain they experienced in the respirator. Those who were unable to tolerate much deprivation believed themselves able to stand pain especially well and did not believe that the parental demand for self-control in their youth, control in the expression of suffering with pain, had been stringent. On the other hand, those persons who were able to stand sensory depriviation well did not think they could stand pain well and believed that the parental demand for self-control in the expression of suffering with pain had been unduly stringent.

Thus it appears that he who is least susceptible to deprivation is most susceptible to pain and that a pain or ache has a different value for a person according to his tendency to reduce the perceived intensity of sensation and his associated personality traits. To examine this generality further we decided to seek from the *Ss* who had

[13] Philip Solomon, P. H. Leiderman, J. H. Mendelson, and Donald Wexler, Sensory deprivation, a review, *Amer. J. Psychiat.*, 114, 1957, 357-363.

undergone deprivation subjective quantitative estimates of the pain they shared in common—the muscular pain associated with being confined in a respirator.

Each of 8 *Ss* was asked what proportion of his earnings from the experiment he would forego could he be relieved of the pain when required to repeat the rest of the experiment. Of the 4 *Ss* who stayed longest in the respirator one offered to take off 100% of his earnings for backache. Another offered 50%, a third 40%, and a fourth said that he had experienced mere discomfort rather than pain and that this had not been an important factor in his thinking.

In contrast to these offers, of the 4 *Ss* who stayed in the respirator the shortest period, one said he would never undertake the experience again for any amount of money, that the experience was infinitely worse than the most agonizing pain he could imagine. Two, who said they had left the respirator because of the pain,

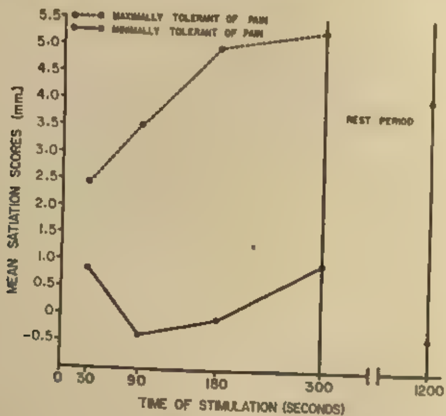


FIG. 1. DECREASE IN SIZE OF TACTUAL TEST-OBJECT FOR GROUPS WHO ARE MINIMALLY AND MAXIMALLY TOLERANT OF PAIN

would not forego any earnings to be relieved of it, and one agreed to forego 25% of his earnings.

It is clear that the absence of pain is more greatly valued by those *Ss* who tolerated sensory deprivation best; and they were the *Ss* who tend to be Non-reducers, who had their experience of pain undiluted. Implicit in these findings is the fact that the non-tolerators of deprivation are much more troubled by other aspects of the experience than by the pain, aspects that constitute the characteristic nature of the stress of sensory deprivation. Perhaps they, nevertheless, complain of pain because our culture regards pain as a proper signal of which to take notice and our language is rich in words for its description. These last two factors are absent, of course, in the stress of sensory deprivation.

Indeed, since pain is sensory, the Reducers might even need the pain of which they complain in the respirator, for it diminishes their sensory starvation and may protect them from hallucinatory substitutes. Two *Ss* in the group exposed to the most stringent conditions of deprivation had hallucinations, and they were the only *Ss* who did not report experiencing any pain in the respirator. This finding, therefore,

supports the hypothesis that pain, in reducing sensory starvation, also reduces the need to produce hallucinatory experience.

Clinical Medicine and Satiability

Clinical medicine provides us with many instances of behavior under pain which could be well explained by a theory of satiability. For example, Dr. Travell of Cornell has reported to me that patients do not complain of pain if an injection of novacaine is made while an area is painful, that is to say, when, according to our theory, some satiation has taken place; yet plenty of complaints are made if such an injection is given in an area that is no longer painful. Dr. Gray, of the Harvard Medical School, has similarly reported that electrical treatment in a new area of the body results in the patient's thinking it is more painful than in the previous area

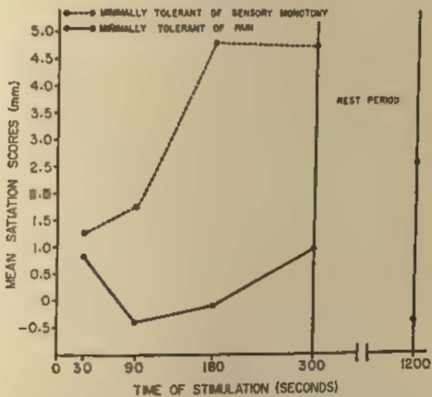


Fig. 2. Decrease in Size of Tactual Test-Object for Groups Who Are Minimally Tolerant (1) of Sensory Deprivation and (2) of Pain

treated. In the new area, according to our theory, he has had no satiation for his pain, and therefore finds it greater. Every good nurse knows, moreover, that she can give injections, without causing pain, by counter-irritation—by firmly pinching the area before she inserts the needle.

Time and Satiability

Besides satiability, there is another perceptual characteristic that changes after a pre-frontal lobotomy and in respect of which suggestive differences exist in contrasting *Ss* who are most and least tolerant of experimental pain. This is, moreover, again a difference that is reversed for those who who are most and least tolerant of deprivation. It is a difference in the perception of time. A minute is experientially lengthened after a pre-frontal lobotomy and, as a corollary, the time spent over a task tends to be underestimated. This altered perception of time is also characteristic of the good tolerator of pain and of the poor tolerator of deprivation.

The mean difference between the good and poor tolerators of pain in the reproduction of 60 sec. is 7.8 ($SE = 4.3$) sec. and between the good and poor tolerators of deprivation it is -6.6 ($SE = 7.1$) sec. If the scores of the poor tolerators of pain and the good tolerators of deprivation are combined and contrasted with the rest of the group, a difference of this size in this direction reaches the 5% level of confidence.

Everyone experiences periods full to the brim and periods curiously empty. The full periods appear to fly past when we are in them, but, when we look back at them, we feel they surely must have taken longer than they actually did. The empty periods, on the other hand, appear to be passing slowly while we are in them, but retrospectively, they seem shortened. It might, therefore, be predicted that the Reducer should experience time which is relatively empty for him, because of his tendency to reduce his perception of the environment and judge it as passing more slowly than does the Non-Reducer. This difference is, indeed, just what we have found.

Another characteristic of the good tolerator of pain and the poor tolerator of deprivation, one that parallels the change after pre-frontal lobotomy, is that such an *S* is more interested in, least worried about, and most happy about the present in contrast to the past and future. Details have been presented elsewhere as to how this change in the perception of the passage of time might contribute to tolerance for pain, as would also a loss of preoccupation with the past and future.[14] Such a patient would be dealing only with present pain and in a more contracted form.

Brief reference may be made to two studies in progress which amplify the findings reported here. First, children's characteristic behavior with pain and their intolerance of perceptual monotony and confinement have caused us to suppose that a tendency to reduce the perceived intensity of sensation might be a characteristic of childhood. Our pilot study suggests that such is the case. Secondly, the results of our preliminary study of 25 alcoholics, who were contrasted with 50 non-alcoholic patients, indicate that there is among the alcoholics a predominance of the personality type of the Non-reducer—of those who suffer much from the insults of the perceptual environment. This relation, taken in conjunction with the fact that alcohol was one of the first anesthetics to be used, suggests that alcohol might cause an increase in the tendency to 'reduce.' We have now found, in a pilot study with 28 members of a hospital staff, that, under the influence of alcohol, the Non-reducers tend to become Reducers.

Discussion and Conclusions

The results of the experiment support these hypotheses: (1) The aspects of personality changed by pre-frontal lobotomy differentiate those who can tolerate pain well from those who suffer greatly with pain. (2) Individual differences in tolerance for pain are paralleled by differences in

[14] Petrie, *Psychopharmacology*, 1958, 99-115; *Proc. 14th int. Cong. Psychol.*, 1954, 167.

perception, especially in the tendency subjectively to reduce the intensity of sensation and in the perception of time. (3) Individual differences in the tolerance for sensory deprivation are also paralleled by these differces in perception, but the direction of the difference is reversed.

Satiability may contribute to each of these tolerances. Indeed, the tendency to reduce the perceived intensity of sensation may be in part the mechanism of tolerance for pain, in that an intermittent bigger wave of pain causes subsequent pain to be perceived as less intense. The tendency to reduce could also be in part the mechanism of the intolerance of sensory deprivation in that it would cause the limited stimulation available to be perceived as even less intense. The sensations of a Reducer would be more diminished, as a result of his previous sensations, than would those of a Non-reducer. One may think of the Reducer as being, in his day-to-day life, subjected to some sensory deprivation: the greater his tendency to reduce, the greater his deprivation. The results of our experiment suggest that the intensity of his sensory experience (including pain) is mitigated because of his tendency to reduce—his tendency to diminish the apparent size of the object. This person is intolerant of further deprivation. At the other end of the spectrum are those whose incoming perceptions are least diminished by previous perception, who have their pain and other sensations 'neat' and undiluted, in whom pain is cumulative, and who more readily tolerate sensory deprivation. Thus different kinds of resistance appear to be needed for tolerating the stress of pain and the stress of sensory deprivation.

Experimental and actual sensory deprivation is never complete and is, indeed, at times a consequence of monotonous stimulation. Just as contrast and change are the conditions of attention, so monotony is in fact psychologically the equivalent of diminution in sensory input. We know that the process of habituation on the psychological level is accompanied by the nervous system's cutting off monotonous stimulation. For example, touch fibers adapt rapidly and respond only to change of pressure. Pain is, however, an exception in that it is not quickly blocked out[15] and under ordinary conditions complete habituation for it rarely occurs.[16] In sensory monotony or deprivation the non-adaptation to any pain that may be present results in a diminution of the sensory starvation: the pain provides 'sensory nourishment.' He who in sensory deprivation is starved of sensation may find that pain, as a sensation, constitutes an alleviation of his stress.

---

[15] E. D. Adrian, *Basis of Sensation*, 1928, 101.

[16] K. M. Dallenbach, Pain: History and present status, this JOURNAL, 52, 1939, 331-347; L. J. Stone and K. M. Dallenbach, Adaptation to the pain of radiant heat, this JOURNAL, 46, 1934, 229-242.

This view gains support from our finding that those most tolerant of deprivation estimate the pain they experience under these conditions as more intense than the pain experienced by those least able to tolerate deprivation. It may well be that the more intense experience of pain is also contributory to the ability of Non-reducers to tolerate experimental deprivation. The fact that only those without pain experienced hallucinations is surely most relevant.

Findings that animals brought up in a restricted sensory and social environment behave as though some forms of pains are not noxious lend further support to this view. For such animals, the normal avoidance of pain is absent. A chimpanzee, whose tactual, kinesthetic, and manipulative experience had been restricted during the first 31 mo. of its rearing, on being subsequently pricked with a pin often responded by panting as chimpanzees do when they are being tickled and enjoying the stimulation.[17] A Scotch terrier who has been reared under conditions of extreme sensory deprivation will put his nose into the flame of a lighted match. This type of behavior does not occur for animals reared normally.[18] It is not impossible that such findings on the attractiveness of pain in sensory starvation may also turn out to have some relation to the origin of masochism in man. The main import of this paper is, however, the manner in which differences in perception are related to variation in the tolerance of pain and the tolerance of sensory deprivation.

---

[17] H. W. Nissen, K. L. Chow, and Josephine Semmes, Effects of restricted opportunity for tactual, kinesthetic, and manipulative experience on behavior of chimpanzee, this JOURNAL, 64, 1951, 485-507.

[18] Ronald Melsack and T. H. Scott, The effect of early experience on the response to pain, *J. comp. physiol. Psychol.*, 50, 1957, 155-161.

THE EFFECT OF PRIOR LEARNING OF SYMBOLS ON PERFORMANCE IN REASONING

By ELI SALTZ, Wayne State University, and SLATER E. NEWMAN, North Carolina State College

To date little is known about the ways in which words influence human behavior. Research using words has concentrated on the learning of words as stimulus-objects and as responses to other words, but has paid relatively little attention to the mechanisms involved in the use of words to control and direct other performance.

One aspect of the effect of verbal symbols on subsequent performance was investigated by Gagne and Baker who taught *Ss* a meaningless name for each of four stimulus-lights.[1] Next, the same stimulus-lights were used as stimuli in a switch-pressing task and the *Ss* were taught which of four switches to press upon presentation of each light. The results indicated that learning verbal symbols as names to the stimulus-lights facilitated subsequent learning of the pressing-responses to these lights. The principal effect of the learning of symbols on learning the subsequent task of pressing the switch was to eliminate incorrect responses, not to speed them. Gagne and Baker argued that learning names reduced the psychological similarity among the lights, making them more discriminable (*i.e.* learning of names resulted in stimulus-predifferentiation). Results of subsequent research on similar tasks have tended to be in agreement with the Gagne and Baker findings, though various theoretical formulations have been offered to explain the results.[2]

Recently, Stolurow and his associates have attempted to extend Gagne and Baker's notion of stimulus-predifferentiation to a more complex situation which we shall call a task of 'decision-reasoning.'[3] The difference between the Gagne-Baker task and

---

* Received for publication May 22, 1958. The experiments reported in this paper were carried out at Chanute Air Force Base, under ARDC Project # 7709, Task 37304.

[1] R. N. Gagne and K. E. Baker, Stimulus predifferentiation as a factor in transfer of training, *J. exp. Psychol.*, 40, 1950, 439-451.

[2] K. E. Baker and R. C. Wylie, Transfer of verbal training to a motor task, *J. exp. Psychol.*, 40, 1950, 632-638; W. F. Battig, Transfer from verbal pre-training to motor performance as a function of motor task complexity, *ibid.*, 51, 1956, 371-378; G. N. Cantor, Amount of pretraining as a factor in stimulus predifferentiation and performance set, *ibid.*, 50, 1955, 180-184; A. E. Goss, Transfer as a function of the type and amount of preliminary experience with task stimuli, *ibid.*, 46, 1953, 419-428; J. S. Robinson, The effect of learning verbal labels for stimuli on their later discrimination, *ibid.*, 49, 1955, 112-114; I. L. Rossman and A. E. Goss, The acquired distinctiveness of cues: The role of discriminative verbal responses in facilitating the acquisition of discriminative motor responses, *ibid.*, 42, 1951, 173-182.

[3] L. M. Stolurow, Mechanical disassembly-assembly as a function of demonstration and part-name learning, University of Illinois Memorandum Report C-8, August, 1953, 1-45; R. H. Burros, L. M. Stolurow, and M. L. Gardner, The relationship of mechanical assembly time to the relevant complexity of the photograph studied, University of Illinois Memorandum Report C-22, September, 1954, 1-19.

Stolurow's can be schematized as follows: In the Gagne-Baker task, the learning of names was followed by rote learning, in which particular lights were arbitrarily paired with particular switches. In this latter task, learning the correct switch to press for a given light depended on memory of the correct *S-R* relation in a previous trial. In Stolurow's task, on the other hand, learning of names was followed by a task in which *S*s could reason what was the correct response from cues in the situation. In one study, Stolurow required *S*s to assemble the components of a machine-gun after learning names for these components.[4] At each step in the assembly, the *S*s had to make a *decision* concerning which component to select next; and these decisions could be determined by cues provided by the components to be assembled. For example, a nut of a given size had to be assembled to a bolt of a corresponding size.

Contrary to Stolurow's expectations, prior learning of names did not facilitate decision-reasoning. In fact, there was some indication of a detrimental effect, though the results were not conclusive in this respect. The first of the Stolurow studies involved a single trial on an assembly task and used speed of assembly as the criterion measure.[5] Results indicated a small, statistically unreliable detriment to assembly performance as a result of prior learning of the names of the components to be assembled.

Stolurow's second study involved four trials and measured performance by number of errors.[6] Each trial was, however, a complex unit involving both disassembly and assembly. Thus, both reasoning and memory were involved. Results of this study indicated that there were detrimental effects, due to learning of names, on the first trial; there was some tendency toward facilitation on subsequent trials. None of these effects was statistically reliable.

It is quite possible, of course, that prior learning of symbols has no effect on the type of task studied by Stolurow. On the other, it is possible that the procedures used by Stolurow and his associates obsecured the relationships which exist.

The first experiment reported in the present paper was designated to investigate the relationship between prior learning of names and decision-reasoning under conditions which eliminated the contaminating influence due to memory of correct responses on the previous assemblies. The *S*s were permitted only one assembly-trial. In addition, since Batig reports that increased motor-complexity reduces the effects of prior learning of symbols, motor-complexity was reduced by requiring *S* to indicate the *order* in which the components of the pressure regulator fitted together.[7] *S*s were not required to execute the acts involved in actually fitting the components together. The task was such that each component provided cues as to its relation with other components. For example, one component was so grooved that only one other component could fit into it. Some cues were subtle; others less so (*e.g.* nut and bolt).

---

[4] Stolurow, *op. cit.*, 1-45.
[5] Burros, Stolurow, and Gardner, *op. cit.*, 1-19.
[6] Stolurow, *op. cit.*, 7-11.
[7] Battig, *op. cit.*, 371-378.

The results obtained in Experiment I indicated that there was a significant relationship between prior learning of symbols and the decisions made. This relationship, however, proved to be non-linear. Therefore, Experiments II and III were performed to study the phenomenon more fully. In Experiment II, the possibility was tested that the learning of symbols affects decision-making because the learning permits familiarization with the components, and that learning of symbols *per se* is irrelevant. Experiment III was performed to determine if the symbols used had any intrinsic value as a cue, either facilitating or detrimental to the performance.

Experiment I

Method: Criterion. The task used as a criterion was so chosen that it involved only the decision-making aspects of assembling the regulator; the manipulative aspects were eliminated. Each *S* was given the shell of the regulator, and the 15 components that fitted into it. These components were arranged in a predetermined unsystematic order on a display board. *S*'s task was to indicate, in the proper sequence, which component fitted into the shell first, which fitted into the shell next, and so forth. *S* could pick up the components one at a time to examine them before making a choice. Upon making a decision, *S* removed the component from the display board and placed it in its proper sequence in front of the shell (*i.e.* the first component closest to the shell, second component behind the first, etc.). After *S* made a choice, *E* told him whether it was right or wrong. If *S* chose correctly, he went on to his next choice. If he chose a wrong component, *S* was so informed. He then returned it to the board and selected a different one. *S*'s score was the number of incorerct decisions he made in choosing components. No time measures were taken and *S* was not deterred from thinking through his decisions.

Symbol-learning. The learning of symbols was conducted as a paired-associates task in which the *actual name* of the component was in the stimulus-position and a picture of the component was in the response-position of a 5 × 8-in. card. Training- and test-trials were alternated. On the training-trials, the symbols and pictures were presented together for 3 sec. On test-trials, *S* was shown a name on the left hand side of a 5 × 8-in. card and five pictures on the right hand side; *S* indicated which of the five pictures went with the name. *S* had 3 sec. in which to respond. The order of pairs was altered from trial to trial and the four alternative parts for any given symbol were altered from test-trial to test-trial. The *Ss* were brought to one of three criteria of learning: (a) no learning, (b) 12 of the 15 possible correct responses on a single trial, or (c) 3 consecutive perfect trials.

Procedure. The *Ss* were individually tested. All *Ss* were shown the board displaying the parts at the beginning of testing. To determine how many names *S* knew prior to learning of the symbols, the names of the parts were read to him in a random order and he was asked to point to the part he thought belonged with each name. A noncorrecting procedure was used. *S* next learned the symbols and then, finally worked on the criterional task.

Subjects. The *Ss* were 36 airmen awaiting assignment to Air Force technical schools. The experiment was initially conducted with 18 men, randomly assigned,

6 men in each of the three cells. The results approached, but did not reach, statistical significance; and so the entire experiment was replicatd with an additional six men per cell. This permitted use of treatment-×-replication analysis of variance.

Results. In Experiment I, small amounts of learning of the symbols facilitated the making of decisions; large amounts of learning, on the other hand, produced marked decrements. As can be seen in Table I, the two replications of Experiment I produced almost identical results; the stability of these data appears to be much greater than is usually found in psychological research. This is particularly striking, since each mean is based on the scores of only six *S*s. The *F* for number of incorrect decisions over the three degrees of learning of the symbols is significant at well

TABLE I
MEAN NUMBER OF ERRORS IN ASSEMBLY AS A FUNCTION OF PRIOR LEARNING OF SYMBOLS

Replications	Amount learning		
	None	Low criterion	High criterion
1	21.67	16.33	28.67
2	21.33	16.17	28.17

beyond the 1% level. Both the facilitation of assembly with small amounts of learning and the decrement with greater amounts of learning are significant beyond the 2% level. Thus, those results for Experiment I which involve greater learning of symbols sustain the trends reported in the Stolurow studies, both of which involved great amounts of such learning.

EXPERIMENT II

It is possible that the results of Experiment I were not a function of learning of the symbols *per se*, but were due to familiarization with the components that occurred during the learning. This possibility was investigated in Experiment II by substituting familiarization, in place of the learning of symbols, prior to performance of the criterional task.

Method. The criterion was identical with that of Experiment I. The material used to produce familiarity was identical with the material to be learned in Experiment I with one crucial modification: the names of the components were covered over so that the *S*s who were permitted to inspect the pictures did not learn the names. The three experimental groups differd in degree of familiarization: Group A was not shown the pictures of the components; they served without familiarization. Group B, low familiarization, was presented with the pictures at a 3-sec. rate for as many trials as the median number of presentations for the low-criterion group in Experiment I (*i.e.* one trial). Group C, high familiarization, was presented the pictures of the components at a 3-sec. rate for as many trials as the median number of pres-

entations for the high-criterion group in Experiment I (*i.e.* 10 trials). Those *Ss* presented the pictures were instructed to study them carefully. In all other respects, the instructions and procedures were identical with those of Experiment I.

Subjects. In Experiment I, the *Ss* were airmen awaiting the start of training in an Air Force technical school. Such *Ss* were not available for Experiment II. Instead, men who had already started training were used. While the intelligence distributions were approximately equal for the *Ss* in the two experiments, it is possible that some men of lower ability or lower motivation were eliminated from the population used in this Experiment. Thirty *Ss* were randomly assigned to the three conditions; 10 in each group.

Results. The results in Experiment I appear to be a direct result of the learning of symbols and not merely due to familiarization with the components of the pressure regulator. This is indicated by the fact that familiarization alone in Experiment II did not significantly influence the decisions made. The group means for Experiment II were 17.7, 20.7, and 17.7 incorrect decisions, respectively, for the groups with no low, and high familiarization. An analysis of variance yielded an $F < 1.00$ over these mean values.

The difference between the decisions made after learning the symbols and after familiarization was evaluated by means of a trend analysis. To perform the trend analysis, the two replications of Experiment I were combined. Then a 2×3 analysis of variance was performed comparing the effect of symbol learning with the effect of familiarization over the three levels of familiarization and learning of symbols. The interaction between levels and familiarization was significantly less than between levels and learning of symbols—beyond the 2.5% level.

Experiment III

There is some possibility that the symbols used in Experiment I may have been misleading.[8] Two examples will illustrate the form this might have taken: (a) the customary order of assembly for a nut and bolt might be *bolt* followed by *nut*, whereas the reverse order may have been demanded in the experimental task; (b) the term *nut* might refer to an object that *S* customarily called a bolt. A careful examination of the names of the components of the pressure regulator did not reveal any obvious false cues. When the names were read to the *Ss* before they were learned in Experiment I, the *Ss* were asked to point out the component referred to by each name; the median number of correct guesses was only 6 of a possible 15. The incorrect choices showed no systematic pattern of false choices over

[8] The writers are indebted to Dr. A. A. Lumsdaine for suggesting Experiment III

all *Ss*. This does not, however, eliminate the possibility that individual *Ss* were responding incorrectly in idiosyncratically regular manners. Consequently, in Experiment III, the *Ss* were asked to indicate the order of assembly using the names alone, with no previous learning of the symbols.

Method. The materials for this study were the 15 names of the components of a hydraulic regulator. These names were identical with those used in Experiment I. Each name was typed on a separate 3 × 5-in. card. The *Ss* were not shown the components or pictures of the components named on these cards. Each *S* was randomly assigned to one of two groups. These groups differed only in the amount of familiarization with the names. Upon entering the experimental situation, each *S* in Group 1 was shown each card once, at a 3-sec. rate, and was asked to read the name on each card. Each *S* in Group 2 was paced through the set of names five times, also at a 3-sec. rate. The order of names during these five presentations was identical with the order used during the learning of the names in Experiment I. After familiarization, the names of the components were spread out before *S* in the same pattern as was used to display the actual components in Experiments I and II. *S* was shown the shell of the pressure regulator and was told that a series of components fitted into the shell in only one order. *S* was instructed to put the names of the components in the same order that the actual components should be fitted into the shell of the regulator. *S* was first to choose the name representing the component that fitted into the shell first; then the name representing the component that fitted into the shell next, etc. A corrective procedure was used and *S* continued to choose until he made a correct response at every point in the sequence. If a name was incorrectly chosen at a given point, that card was turned over, that it could not be chosen again until the correct choice was made. As in Experiments I and II, *S* was permitted to 'assemble' the regulator only once.

Subjects. Ss were 26 airmen awaiting training in an Air Force technical training school; 14 *Ss* were randomly assigned to Group 1, 12 *Ss* to Group 2. The distribution of aptitude scores was comparable to that for *Ss* in the first two experiments.

Results. The results of Experiment III indicate that the symbols *per se* had a small but significant facilitating effect on the making of decisions. If *Ss* had made all their responses in a perfectly random fashion, they would have been expected by chance to make an average of 52.5 errors in ordering the 15 names. The mean number of errors for Group 1 was 42.49, for Group 2, 43.50. Both these means differ from the hypothetical mean of 52.5 beyond the 1% level of confidence. The mean number of errors for Group 1 does not, however, differ significantly from the mean number of errors for Group 2, $t < 1.00$.

Discussion

The results of the three studies reported in the present paper can be summarized as follows. In a task such as the one used in these experiments, in which prior practice on the task is eliminated and in which motor-

complexity is reduced, prior learning of the names of the components affects decisions. The effect appears to be a non-linear one: small amounts of learning of the symbols are facilitating; large amounts are detrimental. Apparently, the results are not an artifact due to familiarization gained with the components to be assembled.

It appears reasonable to suppose that the facilitation with small amounts of learning is due to the value of the symbols as cues, such as found in Experiment III. Had nonsense-syllables been used in place of the actual names of the components, this facilitation might not have resulted. The decrement in performance with greater degrees of learning is more difficult to understand. It is interesting to note that knowledge of names was accompanied by detriments to the making of decisions *only* in the situation in which the names were learned to a high criterion. Merely being given the names, as in Experiment III, did not hinder performance; this was true even for the group, in Experiment III, familiarized with names. This argues that the factors producing the decrement in performance result from the learning of a name associated with a specific component. Two possibilities are suggested: (a) The learning of symbols may lead to *avoiding* the correct decision, *e.g.* errors during learning of symbols may result in the development of affective reactions which lead to the avoidance of the relevant referent objects. Some support for this suggestion is found in the analysis of the error data in Experiment I. The design of the study, however, precluded a definitive test of this possibility. (b) Symbol-learning may lead to *making* an incorrect decision, *e.g.*, some symbols may be learned so well that they generate inappropriate tendencies to choose the referents of these symbols.

If either of the above suggestions proves correct in further research, it seems likely that the learning of symbols will not invariably result in a decrement in performance on a task like ours. For example, if the learning of a name results in an increase in the tendency for certain components to be selected, performance will be facilitated if these components should correctly be chosen early in assembly.

It is interesting to note that the results of Experiment I, and the tendencies in the two Stolurow studies, are opposite in direction from the results of the Gagne-Baker study, and the many others that were similar to it.[8] In the Gagne-Baker study large amounts of prior learning of symbols *facilitated* subsequent performance. There are many differences, of course, between Experiment I and the Gagne-Baker study. In Experiment I, the decision concerning which component to select next for assembly, was uninfluenced by memory of which component was correct on a previous

[8] Gagne and Baker, *op. cit.*, 439-451; see also Footnote 3

trial. The Gagne-Baker effect, on the other hand, is primarily an effect of prior learning on *new learning*. In Experiment I, the next response to be made can be reasoned from cues in the situation. In the Gagne-Baker study, *S-R* pairs were arbitrarily formed and reasoning is irrelevant.

Consequently, the writers feel that the difference in direction of results between the kinds of task used by Gagne and Baker and by Stolurow may indicate that the distinction between *learning* and *reasoning* is an important one in studies of prior learning of symbols. A deal of support would be lent to this position if it could be shown that prior learning of symbols facilitated learning on Stolurow's type of task.

Another problem which arises is the lack of significant results in Stolurow's studies, contrasted with the significant results in Experiment I in the present paper. The results of the present studies suggest that in both Stolurow's studies the lack of significance may have been the result of using criteria that measured too gross a unit of behavior. The criterion in Stolurow's first study appears to have measured both reasoning and learning aspects of behavior.[10] His trials were extremely complex disassembly-assembly units, so that even the first trial probably reflected some learning. In support of this analysis, an examination of Stolurow's data reveals that the detrimental effect of symbol-learning was largely confined to Trial 1. The criterion utilized in the Burros, Stolurow, and Gardner study involved a time-measure in conjunction with complex motor acts; consequently, here again decision was contaminated by other variables.[11]

There are several additional points worthy of mention in connection with the experiments here reported. First, it should be noted that familiarization as well as the learning of symbols may be differentially related to learning and to decision. Thus, whereas in Experiment II familiarization was shown to have had no effect on the making of decisions. Robinson, using as his criterion a task in which *Ss* learned to discriminate between sets of fingerprints, found that familiarization with the fingerprints was as effective in promoting discriminative learning as learning names for the fingerprints.[12] The *Ss* in both his groups—the familiarized and the symbol-learning—performed better on the task of discriminative learning than the *Ss* who had neither prior familiarization nor prior learning.

A second point is that the particular behavior measured in the present experiments produced remarkably stable results between replications. This suggests that the making of decisions relative to an assembly is a stable unit of behavior.

SUMMARY

Three experiments were conducted to investigate the effects of prior learning of a symbol on decisive reasoning. That is, the making of decisions in which the responses are not determined by memory of what responses were correct on previous trials but must be reasoned out on the basis of relationships among cues in the situation. In Experiment I, small amounts of learning were shown to facilitate subsequent decisions, more perfect

---

[10] Stolurow, *op. cit.*, 1-45.
[11] Burros, Stolurow, and Garner, *op. cit.*, 4-6.
[12] Robinson, *op. cit.*, 112-114.

learning of the symbols being detrimental. In Experiment II, familiarity with the components, without learning their names, was utilized. Such familiarization had no effect on the making of decisions. The results of Experiment III, conducted to determine whether the symbols may have offered misleading cues, indicated that the names did have a facilitating effect on the decision that was greater than chance. The results of Experiments I and II, in the light of contrary findings in previous experiments, suggest that both prior learnings of names and familiarization contribute some share in making decisions that is different from the unpracticed reasoning that also plays a part.

RECOGNITION OF DISTORTED MELODIES

By BENJAMIN W. WHITE, Massachusetts Institute of Technology

Every child, with a phonograph at his disposal, has observed that music, when played very slowly or very fast, is hard to identify, and that when it is played backward it frequently defies recognition. Equally well known is the fact that the transposition of a melody into another key affects its recognizability very little, nor does the instrument on which a melody is played make much difference in the ease with which it is recognized. It was the purpose of this study to measure the effects of various transformations on the recognizability of familiar melodic patterns.

Most work on the recognition of auditory pattern has employed recorded speech as the stimulus-material. Since interest has frequently been centered on the improvement of some device for communication, the distortions experimentally introduced have involved impoverishment of the 'signal' by chopping, filtering, or reducing amplitude; or the addition of extraneous material, such as white noise. There have been comparatively few studies concerned with the effects of transformations of auditory patterns on recognition. A pioneering work of this kind was Werner's study of micromelodies in which the log frequencies of familiar melodies were so divided by a constant that a semitone was only one-sixth its normal size in the conventional musical scale.[1] Such melodies were initially difficult to recognize, but with repeated exposure, were correctly identified and took on a more differentiated character.

The transformations employed in the present experiments are similar in kind to those employed by Werner, although all stay within the conventional musical scale. Some are linear, some non-linear, and all are performed on the same set of 10 well-known melodies.

METHOD

Materials. The 10 songs chosen were:

(1) Bicycle Built for Two (Daisy, Daisy)
(2) Deck the Halls with Boughs of Holly
(3) Home, Sweet Home (There's No Place Like Home)

* Received for publication March 27, 1958. The research reported in this paper was supported jointly by the Army, Air Force, and Navy under contract with the Lincoln Laboratory, Massachusetts Institute of Technology.

[1] Heinz Werner, Ueber Mikromelodik und Mikroharmonik, *Z. Psychol.*, 98, 1926, 74-89.

(4) Jeannie with the Light Brown Hair
(5) Londonderry Air (Oh, Danny Boy)
(6) Old Folks at Home (Way Down upon the Swanee River)
(7) On Top of Old Smoky
(8) Red River Valley
(9) Star Spangled Banner
(10) Yankee Doodle

On the basis of informal initial investigation, it was decided that these ten melodies could be easily identified by the *Ss* if played 'straight.' Only the first 24 notes of each of these melodies were used, thus eliminating melody-length as a basis for identification.

The fact that simple transposition has virtually no effect on the ease with which a melody is recognized suggests that it is perhaps the sequence of intervals between adjacent notes in a melody which carries the information. Accordingly the transformations used in the experiment were effected by performing various operations on the intervals between adjacent notes in each melodic pattern. These intervals were given signed integer values determined by the size of the interval in semitones and the direction of the interval, ascending (+) or descending (−). For example, if the first two notes of a melody were middle C, and E above middle C, the first interval was coded as +4. It should be noted that these integral values are a linear function of log frequency. These sequences of 23 signed values for each melody were then subjected to the operations listed below. The symbol at the right of each operation will be used to designate this particular transformation in later discussion.

(1) Doubling the size of the interval.	$(2i)$
(2) Doubling the size of the interval and decreasing its size by 1, regardless of sign.	$[2i(-)1]$
(3) Increasing the size of the interval by 1, regardless of sign.	$[i(+)1]$
(4) Increasing the size of the interval by 2, regardless of sign.	$[i(+)2]$
(5) Decreasing the size of the interval by 1, regardless of sign.	$[i(-)1]$
(6) Decreasing the size of the interval by 2, regardless of sign.	$[i(-)2]$
(7) Algebraic addition of +1 to the interval.	$(i+1)$
(8) Algebraic subtraction of +1 from the interval.	$(i-1)$
(9) Setting all intervals equal to 1 and maintaining sign.	$i/\|i\|$
(10) Randomly permuting the set of intervals.	*Rand i*
(11) Setting all the note durations equal to a quarter note value, leaving pitch intervals unchanged.	$D=k$
(12) Setting all the pitch intervals equal to zero, leaving the durations unchanged.	$i=O$

In all operations involving subtraction of a constant from an interval (Nos. 2, 5, 6, 7, and 8), an interval was not allowed to change sign but was trapped at a value of zero. All the transformations except No. 11 ($D=k$), left the durations of the notes unchanged. The transformed interval-sequences were translated back to sequences of notes, which were in turn transposed so that all passags had the same lowest note.

The musical scores for these 12 variations on each of the 10 songs, together with the 'straight' versions, making a total of 130 selections, were recorded on magnetic tape in scrambled order with the Laboratory's *Memory Test Computer* as the musical instrument.[2] This high-speed, general-purpose, digital computer is equipped

[2] The author wishes to express his thanks to Carma Forgie for her assistance in programming and recording these melodies.

with an audio amplifier and speaker as an output device which made this method of generating music possible. Although the computer is something less than a thrilling musical instrument, it makes possible a degree of control over note-duration, amplitude and frequency—difficult to obtain with human performance on some musical instruments. (It also afforded the experimenter experience in programming such a computer to generate auditory displays.) There was a 15-sec. pause after each selection, during which a recorded voice announced the number of that item.

Procedure. For each experimental session, the *Ss*, all adults and members of the Laboratory, were given lists of the titles of the melodies numbered from 1 to 10, and answer sheets on which were printed 130 ten-alternative, multiple-choice items. The *Ss*, tested in a group, were simply instructed to listen to the melodic patterns, and, after hearing each selection, to decide which of the 10 songs it was most like, and to record their decisions by marking the appropriate space in the answer sheet. They were asked to make a response to every item. In the first session, which lasted 90 min., the 24-note versions of the melodies were played forward. Nine *Ss* participated in this session. Eight *Ss* participated in the remaining three sessions. The second session was identical to the first except that the tape was played backward. Two additional sessions paralleled the first two except that the tape was edited to present only the first 6 notes of each selection, instead of the first 24 notes used in the first two sessions. In the third session the abbreviated selections were played forward, and in the fourth session, the same tape was played backward. At least one day intervened between sessions.

As a check on the familiarity of the 10 songs, the *Ss* were asked at the outset of the experiment to indicate on the page bearing the 10 titles those with which they were not familiar. Familiarity was defined as knowing the words to the song, or the ability to whistle or sing it. Eight of the nine *Ss* indicated familiarity with all the songs. One *S* indicated lack of familiarity with three by this criterion.

Results

Table I shows the percentage of correct identifications to each of the 13 transformations for each experimental session. That the melodies were reasonably familiar, even when played on a digital computer, is seen by the fact that 85 out of a possible 90 (94%) correct identifications were made when played straight in the 24-note version and in the proper direction. The five incorrect identifications did not arise from a single *S*, but four; nor were they concentrated on a single melody, but four. Thus, although the melodies were not all correctly identified in their undistorted versions, the errors were widely spread over *Ss* and melodies.

Looking at the first column to compare the extent to which the various transformations impaired correct identification, one sees immediately that all of the entries are well above 10%, which would be chance expectancy. Even when all pitch information was removed ($i = 0$), the *Ss* were able

to identify 33% of the items correctly on the basis of the duration alone. Elimination of the duration ($D = k$) was one of the least disrupting variations. In this experiment it was reassuring to find that *S*s did not rely heavily on the rhythmic information to make their identifications, since the remaining transformations affected only the melodic patterning. It should be mentioned that this result is attributable to the particular set of melodies employed. Had the set been more homogeneous in melodic pattern, *S*s undoubtedly would have relied more heavily on rhythm and would accordingly have been more disrupted by its deletion. It just happens to be characteristic of most music in our culture that rhythmic patterns are stereotyped and afford a less adequate basis for identification than melody.

TABLE I

PERCENTAGE OF CORRECT IDENTIFICATIONS OF VARIOUS TRANSFORMATIONS OF 10 FAMILIAR MELODIES, LONG AND SHORT VERSIONS, PLAYED FORWARD AND BACKWARD

Transformation	Forward		Backward	
	first 24 notes	first 6 notes	first 24 notes	first 6 notes
'Straight'	94*	94	35	35
$D=k$	88	60	23	15
$i(+)1$	81	84	29	27
$2i(-)1$	80	81	31	30
$i(+)2$	80	76	21	36
$i(-)1$	79	61	24	26
$2i$	78	79	35	25
$i-1$	71	69	17	20
$i+1$	68	75	31	24
$i/\|i\|$	60	69	25	23
Rand i	52	46	21	23
$i(-)2$	44	46	29	27
$i=0$	33	32	26	23
Average	70	69	27	26

* Percentages in the first column are based upon 90 observations (10 melodies×9 *S*), while those in the remaining columns are based upon 80 observations (10 melodies×8 *S*).

Among the transformations altering only melodic pattern, the five which showed the least effect, and which were very similar in the extent to which they impaired recognizability, are distinguished by the fact that they left the relative sizes of the intervals unchanged as well as the sequence of ups and downs. Transformations which disturbed either of these, $i/|i|$, for example, which preserved the sequence of ups and downs, but sacrificed the information in the relative sizes, had a much more serious effect on recognition. The algebraic addition and subtraction transformations, of

course, had opposite effects on ascending and descending intervals. The $[i(-)2]$ transformation because of the restriction that intervals could not change sign, left many intervals at zero, thus eliminating much pitch information and making this transformation similar in effect to constant frequency $(i = 0)$.

The fact that the random-interval transformation had no more serious effect than $[i(-)2]$ or $(i = 0)$, lends support to the supposition that *Ss* depended heavily on interval-information in identifying melodic patterns. This transformation maintains the set of intervals found in the original melody but throws away their temporal sequence. Thus, if a melody has an ascending interval of five semitones, this interval will probably fall in a different temporal part of the pattern, and will, in all likelihood, be in a different pitch range. That is, a given interval in the original melody may be between the two highest notes early in the sequence, whereas in the transformation it may occur between the two lowest notes and toward the end of the sequence. Despite this, the random interval transformation made the melodies no harder to recognize than the $[i(-)2]$ which preserved most of the sequential pattern, but with intervals reduced in size.

TABLE II
Analysis of Variance of Arcsine Transformation of the Number of *Ss* Correctly Identifying 10 Melodies Under 13 Transformations
(24-Note Version, Played Forward)

Source	*df.*	Estimate of variance	*F*	P
Transformations	12	25.62	12.56	<0.001
Melodies	9	14.28	7.00	<0.001
Residual	108	2.04		

By means of an arcsine transformation of the number of *Ss* correctly identifying each of the 130 items played forward in the 24-note version, a two-way analysis of variance was performed, the results of which are shown in Table II. It is clear that the various transformations differed significantly in the degree to which they affected recognizability, as did the melodies.

Shortening the melodic patterns by a factor of four had comparatively little effect either on the over-all percentage of correct identifications, or on the rank-ordering of the difficulty of the various transformations when they were played forward (see Column 2, Table I). An analysis of variance identical to that reported above was done on the data from this second session with similar results: significant differences between transformations and between melodies.

Playing the tapes backward produced a marked change both in the over-all percentage of correct identifications and in the rank ordering of the difficulty of the various transformations. This was so under both conditions of melody-length. The untransformed melodies played backward were correctly identified 35% of the time. Only one of the forward transformations yielded a lower percentage than this ($i = 0$), which deleted all pitch-information. Although the over-all percentage of correct identifications to the backward version of the melodies was significantly above chance expectancy, an analysis of variance of the responses similar to those described above revealed no significant differences between transformations or between melodies.

There is again little difference in over-all percentage correct between the 6- and 24-note versions played backward, but the rank-ordering of the transformations shows little or no similarity between the two conditions.

It should be mentioned that individual differences between *S*s were marked under all conditions and showed significant consistency from one condition to the next.

Discussion

These experiments confirm the observation that temporal reversal of an auditory pattern makes its correct identification extremely difficult. Other transformations which had shown significant and specific effects on recognition of these melodic patterns played forward, showed no such effects on the temporally reversed patterns. Disruptive though temporal inversion was, *S*s nevertheless identified more reversed melodies correctly than would be expected had they been guessing randomly.

Is such temporal polarity peculiar to auditory patterns, or is it to be found in vision as well? It is certainly prominent in reading. A word spelled backward is hard to recognize, as is a word with its letters scrambled. This is hardly surprising, since reading involves successive fixations in a culturally prescribed direction. A movie played backward, however, is, at some level understandable. People are recognizable as people, and many of their actions are easily identified: walking, eating, etc. When the animated cartoon of Heider and Simmel was run backward, *S*s' descriptions of the action were, in several sections, similar to those given to the forward version.[3] No matter which way it was played, for example, the struggle-sequence was a struggle. Similarly, a phonograph record played backward

[3] Fritz Heider and Marianne Simmel, An experimental study of apparent behavior, this Journal, 57, 1944, 243-259.

retains some characteristics which can easily be identified. It is possible to distinguish a male from a female voice, speech from music, and jazz from philharmonic music. Perhaps it is meaningless to ask if temporal reversal of an auditory pattern is more or less disrupting than temporal reversal of a visual pattern, although parallel experiments in the two modalities with suitably chosen patterns might throw some light on the question.

The other main finding in these experiments is that those transformations are the least disruptive which preserve the relative magnitudes of the intervals between successive notes and do not change the temporal sequence. Furthermore, all such linear transformations affect the recognizability of the melodies to a similar degree. Undoubtedly if the additive or multiplicative constants had been large or, as in Werner's experiment, small fractions, recognition would have been much more difficult. Comparable non-linear transformations might, however, have been still more difficult.

One of the more puzzling findings is that temporal reversal has a more serious effect on recognition than random permutation of the set of musical intervals in a melody. Since *Rand i* throws away so much information, one would expect it to impair recognizability significantly more than temporal reversal. Such is not the case; in fact the difference is in the opposite direction.

The results with the shortened melodies cast doubt on one hypothesis as to the strategy employed in identifying the temporally reversed patterns. This was that the *S*s listening to these patterns wait until the terminal notes, normally the opening notes of the melody, and reverse these before deciding what response to make. Such a procedure should be easier with a shorter pattern, since *S* would not be as uncertain about when the sequence would terminate. With such a strategy, one would predict that short reversed patterns would be easier to identify than the long ones. The results fail to support this prediction as there was virtually no difference in the proportion of correct identifications in the two conditions.

Since familiar melodic patterns were used in this experiment, it is hazardous to look for explanations of the differences between melodies in the extent to which various transformations rendered them difficult to recognize. It is unlikely that the differences are simply a function of 'familiarity,' however defined. It should be noted that several of the transformations employed in this study would have different effects on different melodies. A melody containing large musical intervals is relatively unaltered by a transformation such as $(i + 1)$ or $(i - 1)$, whereas a melody composed of small intervals is greatly affected, particularly if intervals are

not allowed to change sign and have zero as a boundary value. In any study of pattern-recognition where manipulation of the stimulus involves transformation, not just impoverishment through brief exposure or reduced amplitude, there is almost certainly going to be a significant interaction between the transformation and certain characteristics of the pattern that will make recourse to the explanatory concept of familiarity, in the sense of number of prior presentations, suspect. It is possible that familiarity may have to be defined not only by the number, but by the variety of prior presentations.

Summary

*S*s were asked to identify short melodic passages which were various transformations of 10 well-known songs. Linear transformations which involved multiplying the intervals by a positive integer, subtracting or adding an integer were least disruptive in terms of the proportion of such patterns correctly identified. Non-linear transformations were more difficult, and temporal reversal of the melodies still more difficult, equalled in difficulty only by a transformation which eliminated all melodic information and left only the rhythmic pattern intact. Melodies were as easily identified from the first 6 notes as from the first 24, either forward or backward.

DESCRIPTION OF LEARNING TO LEARN IN HUMAN SUBJECTS

By Carl P. Duncan, Northwestern University

Subjects who practice on a series of similar tasks show gradual improvement in performance beyond the anticipated improvement on the particular task, a type of positive transfer which has been called learning how to learn, or learning-set.[1] Although learning to learn is supposed to be a commonplace, there is practically no literature on it in the area of human learning, in contrast to the extensive literature on animals. Descriptions of human learning to learn are usually reported only as ancillary findings, while analytic research has rarely been attempted.[2] It therefore seems worthwhile to describe certain characteristics of human learning to learn. The data to be reported were recorded during one of several training-conditions in an experiment on transfer, but the analyses of them which appear here have not been reported elsewhere.[3]

Method. The details of the method can be obtained from the transfer-experiment. In summary, for one training-condition of interest here (Group 10–10 of the transfer-experiment), the *Ss* learned a series of 10 tasks of paired-associates, each task containing 13 S–R pairs. The responses for all tasks were movements of a lever into any one of 13 slots. The 10 different tasks were provided by preparing 10 sets of forms which served as stimulus-conditions. Each of the 13 forms within a set consisted of elaborations on a 'theme,' with a different theme for every set. Because the theme was different for each set, and because the stimulus-conditions were assigned to slots haphazardly, it is unlikely that transfer among tasks occurred on the basis of specific stimulus- or response-generalization, or stimulus-predifferentiation.

The 10 sets of stimulus-conditions were arranged in 10 completely different orders which were assigned to *Ss* in turn; thus, each of the 10 tasks was made up of all the 10 sets. To prevent serial learning, the 13 stimulus-conditions within a task were presented in 12 different orders. Each condition appeared for 4 sec. Each

* Received for publication April 2, 1958. The author is much indebted to Dr. Bruce O. Bergum for drawing the figures in this paper.

[1] For learning how to learn, see J. A. McGeoch, *The Psychology of Human Learning*, 1942, 401; for learning-set, see H. F. Harlow, The formation of learning sets, *Psychol. Rev.*, 56, 1949, 51-65.

[2] Descriptive studies are summarized in J. A. McGeoch and A. L. Irion, *The Psychology of Human Learning*, 2nd Ed., 1952, 306-309; an analytic study is that by D. R. Meyer and R. C. Miles, Intralist-interlist relations in verbal learning, *J. exp. Psychol.*, 45, 1953, 109-115.

[3] C. P. Duncan, Transfer after training with single versus multiple tasks, *ibid.*, 55, 1958, 63-72.

task was practiced for 20 trials, one task per day over two 5-day weeks. Recording was in terms of correct responses per trial.

The *Ss*—40 men and women, undergraduate students—were paid for their services.

Results. Fig. 1 shows mean number of correct responses per trial, for 20 trials, and for each of the 10 tasks. The points in Fig. 1 are the same as those shown in the top left-hand curve of Fig. 1 in the transfer-experiment.[4] As in all graphs in this paper, the points have been fitted by a smooth curve drawn by visual inspection.

The curve in Fig. 1 shows the typical phenomenon of increasing skill in total performance but of decreasing gains in skill in intertask comparisons.

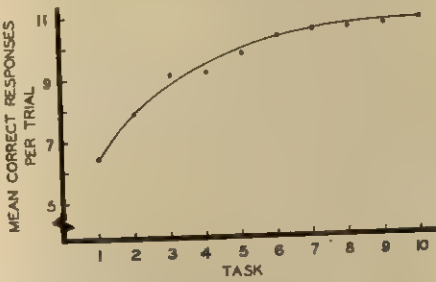


FIG. 1. MEAN PERFORMANCE PER TRIAL OVER ALL 20 TRIALS ON EACH TASK

The over-all improvement was, of course, highly significant; between Task 1 and Task 10 the *t* for related measures was 17.25. Performance on each task was numerically higher than on the preceding task, but most of the total gain was made in the first six tasks. Although 72 hr. elapsed between Tasks 5 and 6 (as compared to the 24-hr. interval separating other adjacent tasks), there was no evidence of forgetting whatever skills enter into learning to learn.

Since there was only a general similarity among the 10 sets of stimulus-conditions, each task had to be learned from scratch; there was only a very slight and irregular improvement from task to task on the first trial of each task. In spite of this, the quantitative gain shown in Fig. 1 is large; performance on Task 10 represents a gain of 65% over performance on Task 1.

Fig. 1 does not show whether learning to learn has differential effects at different stages of intra-task learning. In Fig. 2, performance within each task is broken down into successive blocks of four trials, that is, fifths of

[4] Duncan, *op. cit.*, 66.

practice in this case. Each plotted point shows the mean performance per trial over a block of trials on a particular task.

Fig. 2 shows that learning to learn was most effective fairly early in intratask learning. The gain from Task 1 to Task 10 was largest on the second fifth of practice (Trials 5-8), followed in order of decreasing gains by the third, fourth, first and last fifth of practice. The fact that the least intratask gain was made in the last fifth of practice may be due to the limitation imposed by the asymptote; performance could not exceed 13 correct responses per trial. In later tasks many *Ss* reached this level early in prac-

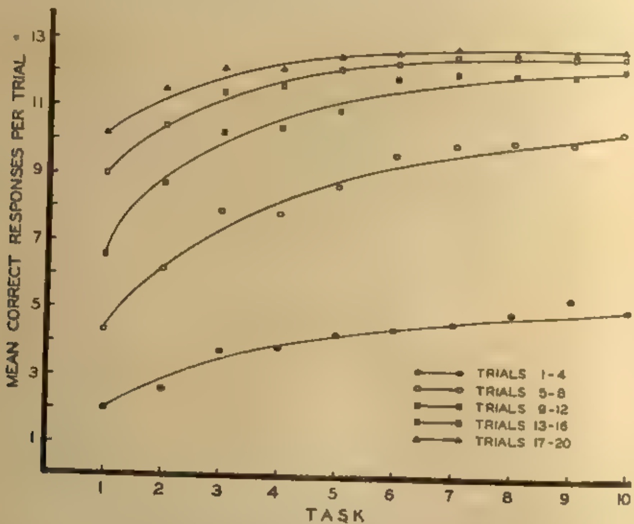


FIG. 2. MEAN PERFORMANCE PER TRIAL ON SUCCESSIVE FIFTHS OF PRACTICE ON EACH TASK

tice; on Task 10 all the *Ss* but one eventually reached 13 correct on later trials.

Another way to look at the same data is to examine the slopes of the learning-curves for individual tasks. Fig. 3 shows intra-task learning-curves for 4 of the 10 tasks. (It should be noted that the four tasks selected for plotting in Fig. 3 are separated by progressively greater numbers of tasks that are not shown.) As can be seen, the learning-curves show progressively steeper slopes. Since some *Ss* were making 13 correct responses on the last few trials even in Task 1, the asymptote of all learning-curves was, again, somewhat restricted by this limiting value. This restriction of the asymptote is quite evident in Task 10, where the means average 12.8

correct responses per trial on the last few trials. Since the question has been raised as to whether slopes or asymptotes of learning-curves are chiefly affected by learning to learn, it may be noted that the present data show that slope may be affected when the asymptote is fixed.[5]

If, in Fig. 3, a percentage-scale is applied to the ordinate, with a mean of 13 correct equal to 100%, it is easy to obtain quantitative indices of how much more rapidly various degrees of mastery were achieved in later than in earlier tasks. For example, 75% mastery was reached on, roughly, Trial 17 in Task 1, whereas in Task 10 this level was reached on approximately

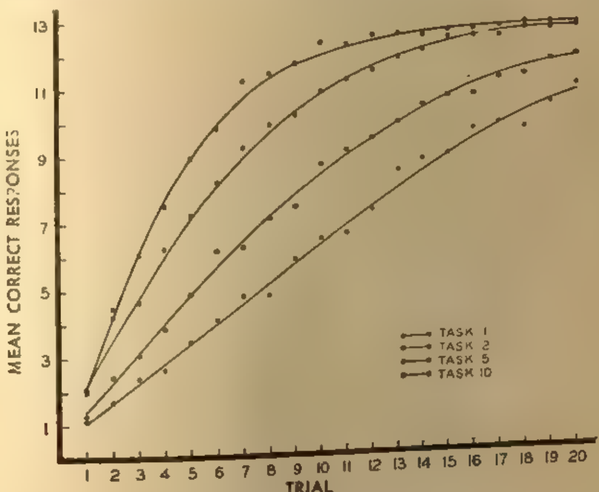


Fig. 3. Learning-Curves for 4 of the 10 Tasks

Trial 7. In Task 1, 90% and 95% mastery were not attained by the end of the 20 trials; in Task 10, these levels were reached just after Trial 8 and just before Trial 11, respectively. These values show that as one goes from low to high levels of mastery the superiority of the later task over the earlier progressively *increases*. This increasingly greater superiority would very probably be larger if it were not for the restriction of the asymptote. It is a function, of course, of the differences in slope noted above.

To demonstrate that as a result of learning to learn, correct responses appear earlier in practice on a task, and the fact that once a response is correct it is less subject to subsequent errors, frequency-distributions were

[5] Meyer and Miles, *op. cit.*, 110.

prepared. The number of responses that were first correctly anticipated out of the total 520 (40 *Ss* × 13 responses) was tallied as well as the number for which the last error was made, at each trial. The frequencies for each of these measures were converted to cumulative percentages and plotted in Fig. 4. (In this figure, there is a point on the curves for last error at Trial 0, indicating the few responses for which no errors were made; these

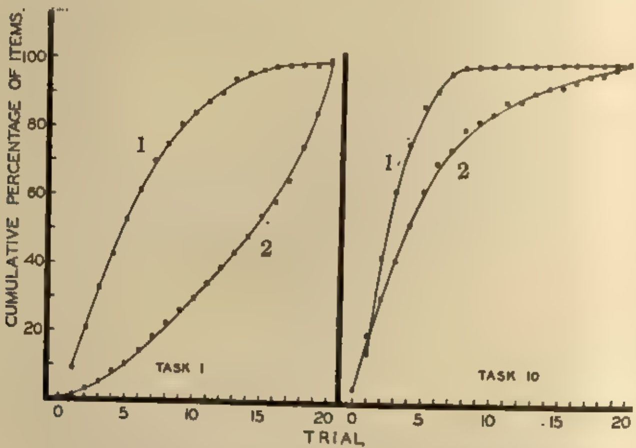


FIG. 4. CUMULATIVE PERCENTAGES OF RESPONSES ON EVERY TRIAL THAT WERE FIRST CORRECTLY ANTICIPATED AND FOR WHICH THE LAST ERROR WAS MADE
1 = First correct anticipation; 2 = Last error

items were included among those which were first correctly anticipated on Trial 1.)

Fig. 4 shows improvement in both measures from Task 1 to Task 10 (the only tasks included in this analysis), but it appears there was greater improvement in the measure of last error. About 75% of the responses had been first correctly anticipated by Trial 8 in Task 1, by Trial 4 in Task 10. For the same percentage of items, the last error was not made until Trial 18 in Task 1, until Trial 7 in Task 10. Other percentages of items give similar results; there was greater improvement in the measure of last error. Thus, a naïve *S* ordinarily gets few responses correct early in practice, and even those are usually unstable in that *S* lapses into errors several times on later trials. As a result of learning to learn, *S* gets more responses correct after fewer stimulus-exposures, and he is much *less* likely to make later errors on those responses. In Hull's terms, learning to learn materially re-

duces the probability that response-oscillation will cause a superthreshold response to fall again below the threshold.[6]

The final analysis of the data is concerned with differences among the *Ss*. Out of the 40 *Ss*, the 10 fastest learners and the 10 slowest learners were selected on the basis of total correct responses on Task 1. Performance of the top and bottom quartiles on Tasks 1 and 10 is plotted in Fig. 5.

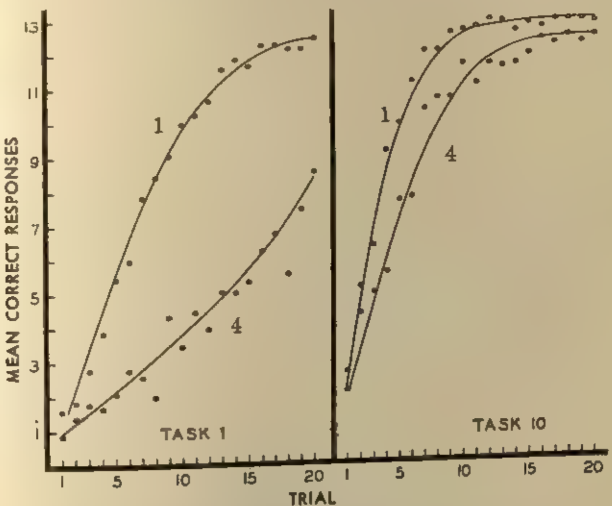


FIG. 5. PERFORMANCE ON TASKS 1 AND 10 OF THE *Ss* IN THE TOP AND BOTTOM QUARTILES ON TASK 1
1 = Top quartile of *Ss*; 4 = Bottom quartile

Fig. 5 shows that both groups of *Ss* improve from Task 1 to Task 10; the improvement of the bottom quartile is striking. On Task 1 the curve for this quartile, although fitted only by visual inspection, is probably not incorrect in suggesting increasing rather than decreasing gains. This gross difference in performance is much reduced on Task 10. (Part of this reduction is probably due to regression.) An analysis of variance for trend of groups over tasks revealed that the interaction of Groups × Tasks was highly significant ($F = 50.44$, 1 and 18 *df.*, $P < 0.01$); this reflects the greater gain made by the bottom quartile. It cannot, of course, be concluded that learning to learn is of more benefit to this quartile because the top quartile was severely restricted by the asymptote. Although the bottom quartile more than doubled its performance from Task 1 (Mean total

[6] C. L. Hull, *Principles of Behavior*, 1943, 332-348.

correct responses = 82.1) to Task 10 (Mean = 196.6), it still did not greatly exceed performance of the top quartile on Task 1 (Mean = 174.9).

Fig. 5 suggests that variance in performance was reduced from Task 1 to Task 10. This is the case if one compares the variances based on all 40 *Ss* ($F = 3.46$, 39 and 39 *df.*, $P < 0.02$), but in the bottom quartile alone the variance significantly increased ($F = 3.59$, 9 and 9 *df.*, $P < 0.02$). Thus, it is quite likely that over-all variance would have increased if the tasks had permitted unlimited performance. The restriction in over-all variance probably accounts in part for the fact that the correlation between Tasks 1 and 10 ($r = 0.54$), although significant at the 1% level, is not especially high.

Summary

Learning to learn by human *S*s was described in detail. Both raw and derived measures showed that learning to learn greatly improved rate of learning (especially on early trials), effectiveness of reinforcement, and freedom from overt oscillation. It was shown that slow-learning *S*s, those in the bottom quartile, benefit enormously from learning to learn, though it was not proved that they benefit more than fast-learning *S*s, those in the top quartile. It appears that the improvement in performance produced by learning to learn is so great that to measure it adequately, the tasks used should have practically no limit on possible achievement.

CONTROLLED FIXATION OF THE STIMULUS-FIGURE IN A STUDY OF AUTONOMOUS CHANGE IN THE MEMORY-TRACE

By CARL P. DUNCAN, Northwestern University

The hypothesis of autonomous change in the memory-trace has not in general been supported, but in many cases the experiments have not yielded either clear-cut or consistent results.[1] Because of this, many attempts have been made to develop better methods in the study of autonomous change. The general method typically used consists of brief inspection by *S* of a stimulus-figure, followed by tests for retention given after different time-intervals. Most attempts to improve method have dealt with the procedures used to measure retention, *e.g.* whether retention should be measured by successive reproductions, by reproductions of independent groups, or by recognition.[2] Little attention has been paid to the initial inspection-period; *S* is usually instructed merely to look at the stimulus-figure for a limited time, presumably in any manner he chooses. This relative lack of control of inspection may be particularly serious in experiments on autonomous change, where interest is often centered on a certain detail of the stimulus, such as a gap or other irregularity. In the absence of specific fixation, or some other control over inspection, there is no way of knowing how the stimulus-object or its details were perceived.

It is the purpose of the present experiment to study a particular autonomous change (closure) in a situation where fixation of the stimulus-object is controlled during the original inspection-period (but, as usual, not during the test for retention) in order to minimize haphazard variations in initial perception of the stimulus-figure. Reduction of variability in the in-

* Received for publication July 28, 1958. This study was supported by Research Grant B-792 from the National Institutes of Health. Dr. Bruce O. Bergum and Mrs. Ruthanne DeWolfe assisted in this study.

[1] Older studies are summarized in J. C. Crumbaugh, Temporal changes in the memory of visually perceived form, this JOURNAL, 67, 1954, 647-658. More recent studies are J. Brown, Distortions in immediate memory, *Quart. J. exp. Psychol.*, 8, 1956, 134-139; J. B. Carlson and C. P. Duncan, A study of autonomous change in the memory-trace by the method of recognition, this JOURNAL, 68, 1955, 280-284; Lawrence Karlin and Glen Brennan, Memory for visual figures by the method of identical stimuli, this JOURNAL, 70, 1957, 248-252; E. L. Walker and Joseph Veroff, Changes in the memory-trace for perceived forms with successive reproductions, this JOURNAL, 69, 1956, 395-402.

[2] Crumbaugh, *op. cit.*, 647-650.

spection of the figure should help to provide a less ambiguous situation within which closure over time, if it occurs, can be measured.

Method

Apparatus. The standard stimulus-figure for all conditions was an outline-circle with a gap. This figure was first drawn in India ink, then photographed and reproduced as a glossy print. In this form the circle was 2⅑₃₂ in. in outside diameter, ⅛ in. in width of outline, with a 4.5 mm. gap (approximately 18°) at the 8 o'clock position. The fixation-point was a small dot, 1/32 in. in diameter, in the center of the circle. Choice of the center of the circle as the fixation-point seemed to be the best compromise between the requirement that the stimulus-figure be fixated, and the custom, in studies of retention of visual form, of avoiding calling *S*'s attention to the particular detail of the figure for which retention will be tested.

The glossy print carrying the standard circle was mounted on cardboard and presented to *S* in a window, 3¾ in. square, cut in a vertical screen of white posterboard, 28 × 37 in., which was fastened to the edge of a table. In this position the fixation-point in the center of the circle was 47 in. from the floor. The circle subtended a visual angle of 1°23′ at *S*'s eye. Under the room-illumination used throughout the experiment, the luminance of the white area of the glossy print bearing the standard stimulus-figure was about 9.1 ml. (measured by the Macbeth Illuminometer).

There were 16 figures (broken circles) for recognition, identical to the standard figure except that the size of the gap varied from 0.5 to 8 mm. in steps of 0.5 mm.; thus, figure No. 9 had a 4.5-mm. gap, the same as the standard figure. Each of the 16 glossy prints bearing a recognition-figure was trimmed to a disk, 3 3/32 in. in diameter, with the broken circle centered within it. The 16 circular prints were mounted edge to edge in the form of a large circle on a sheet of white posterboard (hereafter called the recognition-card), 28 in. square. The 16 prints were so mounted that the gaps within the circles were at 8 o'clock. The prints were lettered *A* to *P* beginning with *A* at the top of the recognition-card and proceeding clockwise. This method of presenting all of the recognition-stimuli simultaneously was used in an attempt to minimize a finding of a previous study, viz. that when the recognition-figures were presented one at a time, *S*'s choice was influenced in large part by the order of presentation.[8]

To minimize biases due to location of the recognition-figures on the recognition-card, three such cards were prepared, with the circular prints mounted in a different, random order on each. Randomness of location was so restricted that circles with gaps of consecutive sizes, and the circles with the largest and smallest gaps were not adjacent.

As a further control over the fixation of the standard stimulus, fixation-time and retention-intervals were varied. There were three fixation-times: 4 sec., 2 min., and 5 min.; and three retention-intervals (test-periods): 0 (immediate test), 1 day, and 1 week. An independent group of *Ss* was assigned to each of the 9 conditions (3 fixation-times × 3 test-periods).

Subjects. The *Ss* were 324 undergraduate men and women, 36 in each of the 9 groups. The *Ss* were assigned to groups in turn. When asked to serve in the ex-

[8] Carlson and Duncan, *op. cit.*, 283.

periment, the *Ss* were informed that "only those with fairly good vision—or good vision corrected with glasses—could serve," hence no *S* had to be rejected for failure to see the fixation-point on the standard-figure from the viewing position of 7 ft.

Procedure. After being seated at the correct viewing distance, *S* was told to look at the small opening in the large white screen which he was facing. Upon a signal, a shutter covering this opening would be raised and he would see a circle with a small dot (the fixation-point) in the center. *S* was told that he was to stare at this dot the moment the shutter opened. (The gap in the standard circle was not mentioned at any time.) He was to continue staring at the dot without moving his eyes or body until the shutter dropped again. It was repeatedly emphasized that although *S* could blink, he must never for a moment take his eyes from the fixation-point or move his head or body. *S* was also told that at some time after this part of the experiment there would be a brief second part (but not what this second part would involve).

At the second session, which followed as quickly as possible for *Ss* in the immediate-test groups, *S* was told that he would be shown a group of circles similar to the one he saw in the first session. He was asked to "look carefully at these circles and select the one which you think corresponds to the original circle." No mention was made of gaps, and nothing was said about fixation. Presenting the recognition-card and giving these instructions took less than 10 sec., but *Ss* were allowed to take their time in choosing a circle from the recognition-card. Most *Ss* took less than 1 min. to make their selection.

The three recognition-cards were used in alternation, so 12 *Ss* in each group were tested on each card.

Results

The score for each *S* is the size of the gap, in millimeters, in the circle *S* chose on the recognition-test. Fig. 1 shows the mean size of gap as a function of retention-interval, with fixation-time as the parameter (each point in the figure based on 36 *Ss*). Fig. 1 also shows a heavy line at 4.5 mm., the size of the gap in the standard circle.

It is obvious from inspection of Fig. 1 that there is no evidence for closure in the original sense of a reduction of the size of the gap over time to values progressively smaller than the initial size. All of the means are numerically above the standard value of 4.5 mm. At the same time, the means appear to vary as functions of both variables, increasing with fixation-time and decreasing with retention-time. These apparent trends are supported, although not strongly, by an analysis of variance. Both variables were significant at the 5% level; the *Fs* were 3.51 for fixation-time, 3.65 for retention-time (each based on 2 and 315 *df.*). The interaction term yielded an *F* less than one. Thus, mean judged size of gap tended to increase with fixation-time to values larger than the standard size, and to decrease from these values over rest, without at any time falling below the standard size.

Since the interaction from the analysis of variance was not significant, groups were combined to yield three grand means on each variable, each mean based on 108 *S*s. By using the within-groups term from the analysis of variance to compute an error-term for *t*-tests, it was found that among the means for fixation-time the 4-sec. group differed significantly from both the 2-min. ($t = 2.09$) and the 5-min. ($t = 2.54$) groups, but the 2-min. group did not differ from the 5-min. group. The *t*-tests on the means for retention-interval showed that the immediate-test group differed

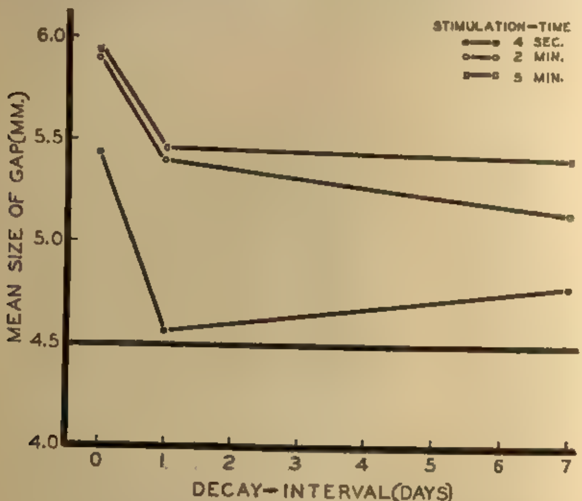


FIG. 1. MEAN JUDGED SIZE OF GAP AS A FUNCTION OF RETENTION-INTERVAL WITH FIXATION-TIME AS THE PARAMETER

significantly from both the 1-day ($t = 2.35$) and the 1-week ($t = 2.44$) groups, but the 1-day group did not differ from the 1-week group.

The results can also be analyzed by comparing each of the means in Fig. 1 with the size of the gap in the standard circle. On the immediate retention-test all three means for fixation-time were significantly larger than 4.5 mm. (*P*-values of 0.01 or less with 35 *df*.). On the 1-day test, only the 2-min. and 5-min. means were significantly above 4.5 ($P = 0.01$ in both cases). After 1 week only the 5-min. mean was significantly larger than the standard ($t = 2.77$, P 0.02); the 2-min. mean did not quite reach significance ($t = 1.97$) with 35 *df*. Although the absolute values of these *t*s must not be taken too literally, since there may be constant errors in this situation, the trends of the *t*-values support the findings of the analysis of variance.

Discussion

It was pointed out earlier that most studies of autonomous change have not yielded unambiguous results. (In fact, in a very recent study James again found, as did Walker and Veroff, that in studying closure, different results can occur merely as a function of the size of the gap in outline-circles.)[4] Unfortunately, the results of the present experiment are also ambiguous; the data neither support nor contradict, in any clear-cut fashion, the hypothesis of closure. It is true that there was no evidence of a narrowing over time of the size of the gap from its initial real size. Mean judged size did decrease during rest, but the fact that all mean judgments were overestimations of the real size prevents identifying this decrease with closure in the original sense.

The consistent overestimation of the size of the gap raises the question as to what factor might have produced such a large constant error. One possibility is that although *S* was required to fixate the center, not the gap, of the standard circle, it is quite likely that during the recognition-test he looked at the gaps in the test-circles. This would mean that every time *S* momentarily fixated the gap in one of the test-circles, he was trying to compare a gap which was immediately present and was being fixated directly, with what he remembered of a gap which was not present and had been perceived indirectly. It would not be surprising if this difficult comparison yielded a constant error, although why the error should have been an overestimation is not obvious. Whatever the source of the overestimation, its effects apparently tend to increase as fixation-time increases.

The judgments of the *S*s may also have been influenced by the method of presenting the recognition-figures. Since the data of an earlier study had been biased by presenting the recognition-figures successively, a procedure for exhibiting them simultaneously was worked out for this experiment.[5] Lovibond, however, quite rightly points out, in a paper published after the present experiment was completed, that any method wherein more than one recognition-figure is presented to the same *S* is likely to be subject to the errors of successive presentation.[6] When *S* has more than one recognition-figure to deal with, there is always the possibility of interaction among the traces of these figures. Since such interaction could occur as easily, perhaps more easily, among recognition-figures that are presented

[4] H. James, Guessing, expectancy and autonomous change, *Quart. J. exp. Psychol.*, 10, 1958, 107-110; Walker and Veroff, *op. cit.*, 397-402.

[5] Carlson and Duncan, *op. cit.*, 283.

[6] S. H. Lovibond, A further test of the hypothesis of autonomous memory trace change, *J. exp. Psychol.*, 55, 1958, 412-415.

simultaneously as among those that are presented successively, this complication may have influenced our results. Lovibond eliminated this problem by presenting only one recognition-figure to each *S* (a procedure which has been used by others but only with retention-intervals of a few seconds) and found no evidence for autonomous change over retention-intervals of 3 min., 1 week, or 2 weeks.

The preceding considerations suggest that in the study of autonomous change in the trace, the method of presenting multiple recognition-figures should be abandoned, but that the method of requiring fixation of the standard figure should be further explored.

Summary

Fixation of the stimulus-figure during the inspection-period was controlled in a study of closure as an autonomous change in the memory-trace. Independent groups of *Ss* fixated an outline-circle with a gap for 4 sec., 2 min., or 5 min., then were tested for retention of the size of the gap immediately, 1 day, or 1 week later. The results were that, at minimal levels of significance, mean judged size of gap increased as a function of fixation-time and decreased over rest. Mean judged size of gap did not however, at any time decrease to values smaller than the real size. Certain methodological problems, which may have contributed to the ambiguity of the results, are pointed out.

SUBLIMINAL STIMULATION IN PROBLEM-SOLVING

By EUGENE O. GERARD, St. Louis University

Of special interest among recent experiments employing 'subliminal stimulation' have been those in which the subliminal presentations prompt the solutions to problems, modify perception, or even suggest specific actions. In each, the problem is whether subliminal stimuli can affect an ongoing perceptual or reasoning process in a direction contrary to that normally taken by these processes.

Kolers found that subliminal presentation of the solution to a problem, involving a common figure in a set of complex geometric figures, resulted in significantly more correct answers than when no subliminal presentation was made.[1] Only correct answers were, however, presented in this experiment; consequently, influencing the subject (*S*) contrary to the normal course of his perceptual and reasoning processes could not be studied. Kolers himself suggested that it would be useful to introduce erroneous solutions.[2] The present experiment replicates the general design of Kolers' study, but with an additional experimental group to which incorrect solutions were presented.

METHOD

Control of subliminal presentation. To have information for solving the problems projected at a level below that producing verbal recognition, Kolers made use of the phenomenon of metacontrast. He describes this as follows:

> If two contours are presented serially and tachistoscopically at appropriate time-values, the first is not reported. If, however, the first is presented alone, it always is reported. The phenomenon thus consists of an apparent inhibition of the first contour by the second.[3]

The term subliminal, as used in this experiment, is specified by the phenomenon of metacontrast. That is, the stimulus is not reported due to the "inhibition of the first contour by the second," rather than due to brevity of duration or low intensity.

Problem-material. The problems in the present experiment were taken from the *Revised Minnesota Paper Form Board*, Series MA (MPFB).[4] Three simple problems

* Received for publication July 29, 1958. The writer is indebted to Dr. Marilyn K. Rigby and Dr. Francis L. Harmon for direction throughout this study.

[1] P. A. Kolers, Subliminal stimulation in problem-solving, this JOURNAL, 70, 1957, 437-441.

[2] Kolers, *op. cit.*, 441.

[3] Kolers, *op. cit.*, 437.

[4] By permission of The Psychological Corporation.

were used for practice and six of the most difficult for the actual testing. They were so arranged that they would be similar in format to those used by Kolers. A composite figure was located at the top center of the field. Across the bottom of the field four (out of a possible five) alternative assemblies of the top figure were placed in a row, and numbered 1, 2, 3, 4; only one of these was a correct reassembly. In the center of the field was a space where first the correct or incorrect solution was projected alone, and this was followed by a grating consisting of 13 vertical and 17 horizontal lines shown as part of the problem-presentation. The correct and incorrect solutions presented subliminally were matched for the number of times they were taken from each position of the four answers.

Apparatus. Kolers used an elaborate tachistoscope to obtain the effect of metacontrast. It was found, however, that comparable effects could be obtained by presenting the problems and solutions from motion picture film. The projector used (Kodak Pageant, Model AV-071) operates at 24 frames per sec. This projector has a shutter which projects each frame twice for 13 m.sec. in alternation with the two interruptions, each of 8 m.sec. Consequently, a single frame is actually projected for a total of 26 m.sec., the screen being dark for the other 16 m.sec. out of each 42 m.sec. cycle.

The first presentation in a given trial was of a solution in the center of the field. This was a single frame. It was followed by 36 frames of the pattern carrying the problem and the grating. Then followed 48 blank frames to allow *S* time for report. If one neglects interruptions during which the pattern does not change, the cycle may be described as: first pattern, 34 m.sec.; pause, 8 m.sec.; second pattern, 1.492 sec.; pause, 8 m.sec.; blank field, 2.008 sec.; pause, 8 m.sec.

The duration of the second pattern, that is, of the problem and the grating, in the present study departed decidedly from that used by Kolers. It is practically five times as long. There were two reasons for this. First, a pilot study indicated that the *S*s, presented with a shorter exposure, reported their answers to be entirely guesses. It was decided in the present study to allow time for more correct answers to develop. Secondly, the *S*s with the shorter duration were frustrated and even hostile when it was impossible to report correct answers. With the longer exposure-time there was noticeably better coöperation and effort to work out the correct answer.

For every *S* the sequence of presentations ran as follows: (1) practice problem No. 1, 25 sec.; (2) practice problems Nos. 2 and 3, 1.5 sec. apiece, with a 2.0-sec. interval between them; (3) experimental session, six consecutive exposures of each of the six problems, a total of 125.5 sec. Every *S* served once under a single experimental condition.

The experiment was run in a small room that contained a table for the projecting equipment and a tape-recorder used in administering the directions. *S*'s chair was located on one side of the table and *E*'s on the other side. The projection screen was 94 in. from the projector and 104 in. from *S*.

The total illuminated field on the screen was 17 × 13 in. The luminosity of the stimulus-field, measured at *S*'s chair by a General Electric Foot-Candle Meter, was 0.2 foot-candles for the blank frames between patterns and 0.5 foot-candles for the pattern-frames (including the blank first-pattern frame for the control group). When the projector was off, there was 0.025 foot-candle of light at *S*'s chair.

Procedure. The directions were on tape. Some sections were taken directly from the MPFB or adapted from it.

After *S* was informed that this would be "a test of rapid perception and rapid reasoning," the first practice-problem was begun. The tape was so timed that, as this problem was being shown, the following instructions were being given.

At the top of the screen there is a pair of triangles, across the bottom four figures numbered 1, 2, 3, 4, and a grating separating the top from the bottom. Now look at the four figures at the bottom. You are to decide which figure shows how the parts at the top can fit together.

The instructions then explained that the test-problems would be of brief duration and that it would be necessary to give some answer to every presentation, whereupon two more practice-problems were shown for the same period of time and with the same intervals that would be used in the test itself. Further instructions explained that each problem would be repeated six times. Questions about the procedure were then answered, but care was taken to avoid information in addition to that included in the general instructions. The experimental session proper was then begun, with the answers being recorded by *E*.

When the testing session was completed, *S* was given a short questionary. Some of the questions were designed to obtain information about factors other than the measured variables that might be influenced by the subliminal presentations. No pertinent information along this line was disclosed. Other questions were designed to determine whether the single-frame solutions were in fact 'subliminal,' *i.e.* whether *Ss* reported seeing them or not. The responses to these questions will be discussed later.

Subjects. The *Ss* were 66 volunteers from the School of Philosophy and Letters. They were seminarians in the course of studies leading to the priesthood in the Jesuit Order. In general, this sample may be considered more homogeneous than a random sample from a general college population in terms of intelligence test-scores, values, and willingness to coöperate. All *Ss* reported approximately normal or fully corrected vision.

The *Ss* were assigned at random to the three experimental conditions. Each condition was assigned 22 *Ss*.

RESULTS

It was of primary importance to determine whether the solution projected for a single frame in Condition A (correct solution presented subliminally) and Condition B (incorrect solution presented subliminally) could validly be classified as subliminal. Whenever *S* indicated in the inquiry at the end of the testing that he had noticed something "not included in the instructions" or "not expected," he was further questioned to determine precisely what it was he had noticed. The answers in all cases referred to irrelevant matters. In not one case out of the 66 did *S* report seeing the single-frame solution. Thus the effectiveness of this method in putting a stimulus outside verbal control was demonstrated.

The responses in the experimental session were scored in two ways: (1)

Number of problems (out of six) correctly answered on each of the six successive presentations of the problem. (2) Number of incorrect answers (*i.e.* the specific incorrect answer that had been projected subliminally in Condition B) for each of the six presentations.

(1) Correct answers. The main findings are presented in Table I. An analysis of variance was applied to the data with the results shown in

TABLE I

MEAN NUMBER OF PROBLEMS CORRECTLY SOLVED, OUT OF SIX, AS A FUNCTION OF CONDITION OF TESTING AND SUCCESSIVE PRESENTATIONS OF EACH PROBLEM

(Each condition includes 22 *S*s)

Conditions	Mean correct solutions	Presentations			
		No.	*M* correct	No.	*M* correct
A (correct solution given subliminally)	13.73	1	1.79	4	2.48
B (incorrect solution given subliminally)	12.50	2	2.11	5	2.73
C (no solution given subliminally)	15.95	3	2.27	6	2.68

TABLE II

ANALYSIS OF VARIANCE OF SCORES (PROBLEMS CORRECTLY SOLVED)

Six presentations to three groups, each consisting of 22 *S*s, tested under different conditions.

Source of variation	Sum of squares	*df.*	Mean square	*F*
Conditions	22.49	2	11.25	7.97*
Presentations	43.02	5	8.60	6.09*
Interaction: C×P	4.15	10	.42	.29
Within groups	533.63	378	1.41	
Total	603.29	395		

* Significant beyond 1% level of confidence.

Table II. A score in each case is the number of problems correctly solved. The *F*s for both conditions and presentations are significant beyond the 1% level; the interaction is not significant. Bartlett's test failed to indicate significant heterogeneity of variance. Tukey's procedure gave a 'significant gap' of 0.708 (5% level).[5] Applied to the means of the conditions (Table I) this indicates a significant superiority of Condition C (control; no solution presented) over both Conditions A and B, and of Condition A over Condition B. These findings, the first of which runs contrary to the orig-

[5] A. L. Edwards, *Statistical Methods for the Behavioral Sciences*, 1954, 331-332.

inal hypothesis that a group receiving correct answers subliminally would do better than a control group, will be discussed later.

The significant F for the six presentations, along with an inspection of the six means, shows a continued improvement through the series of trials. This indicates strongly that more than guesswork was involved in producing correct answers, *i.e.* that the situation did involve a perceptual or reasoning process, which was the purpose of the relatively long exposure-time for the problems.

(2) Incorrect answers. By 'incorrect answers' is meant, as stated earlier, not the total number of all incorrect answers, but the number of times the specific incorrect answer was reported that had been projected subliminally in Condition B. These responses were analyzed in the same factorial design as the correct answers, giving an F of 1.67 for the conditions, an F of 0.73 for the presentations, and an F of 0.55 for the interaction. None of these is significant, indicating that these responses were not influenced by the various experimental conditions, failing to confirm the hypothesis originally proposed.

Discussion

The original hypothesis predicted a positive effect of subliminal presentation of correct answers, and a negative effect of incorrect answers. Condition A should have produced more correct answers than Condition C and Condition B should have produced fewer. The latter of these results was verified; but the control group (Condition C) proved superior to Condition A as well as to Condition B. This finding is in contrast with Kolers' results and warrants examination of some of the possible causative factors.

The most obvious factor would seem to be the presence of unexpected group differences in ability to perform the MPFB problems. In an attempt to control this source of error, the MPFB was administered in its entirety to all the Ss about two weeks subsequent to the experimental sessions. There was no significant difference among the groups, and, in fact, the correlation coefficient within the groups between the experimental scores and the MPFB score was found to be 0.003.

It can be argued that, although subliminal stimulation may have decided influence in guessing or chance situations, its positive influence might not be significant when the perceptual and reasoning processes are presumed to be strongly determined. The superiority of Group C over Groups A and B even suggests that the subliminal presentations interfered with the correct solution of the problems. If, however, this interference was the only result of the subliminal presentations, how can the significant superiority

of Group A over Group B be accounted for? Such a single-factor theory is insufficient to account for the results of the present experiment. At least two factors must be involved.

Summary

The effects of subliminal presentation of a solution together with the supra-liminal presentation of a problem to be solved were investigated in a design similar to that of Kolers. Both the correct solution (Condition A) and the incorrect solution (Condition B), as well as no solution (Condition C), were presented to separate groups of *S*s.

Metacontrast was used to put the cue of the solutions below the level of verbal recognition. It was produced by putting the problems and solutions on motion picture film. The results indicated that with this technique the *S*s did not report seeing the masked cue.

The number of correct responses was found to be significantly fewer following the subliminal presentation of correct and incorrect solutions than in the control condition. This was contrary to the original hypothesis.

RESPONSE-TENDENCIES IN ATTEMPTS TO GENERATE RANDOM BINARY SERIES

By PAUL BAKAN, Michigan State University

This paper is concerned with the problem of response-patterning when *Ss* attempt to generate a random binary series. It has been suggested by Reichenbach that a person, not trained in the theory of probability, if asked to construct a binary series that seemed to him to be random, would construct a series with too many alternations, *i.e.* too many *runs*.[1] The number of runs in a sequence is defined as the number of unlike neighbors in the series plus one.[2] Thus in the series H/T/HH/T/H there are five runs; in the series TTT/HHH there are only two runs.

An empirical evaluation of the hypothesis of more runs than chance expectation, carried out by Ross, did not confirm the hypothesis.[3] In fact it was found that there was a slight but statistically insignificant tendency in the direction of fewer runs than chance prediction. Ross concludes that *Ss* who are instructed to construct a random series give a fairly good approximation to the expected number of alternations.

It should be pointed out that this conclusion is based on an analysis of the average group tendency. Neither Reichenbach nor Ross considers individual differences between *Ss* in the nature of the response-sequences generated. An analysis of individual differences would seem to be important if inferences are to be made about 'man' as a random generator. A failure to find more than a chance number of runs for the group does not necessarily imply that man is not a systematically biased randomizer. It may mean that in a group of *Ss* some have a consistent bias in one direction and others have a consistent bias in the opposite direction and the biases cancel each other out. A closer analysis of such data in terms of individual differences and the reliability of individual differences appears to be in order.

The purpose of the present study then is two-fold: (1) to determine whether individual *Ss* are consistent in their patterning of a binary series when instructed to generate a random series; and (2) to allow for another

* Received for publication August 25, 1958. The author is indebted to Ralph Stewart for assistance in the analysis of the data.

[1] Hans Reichenbach, *The Theory of Probability*, 2nd. ed., 1949, 153.

[2] William Feller, *An Introduction to Probability Theory and Its Applications*, 2nd. ed., 1, 1957, 40.

[3] B. M. Ross, Randomization of a binary series, this JOURNAL, 68, 1955, 136-138.

test of the hypothesis that an *S* will generate more runs than chance predicts under conditions somewhat different from those reported by Ross.[4]

Procedure. Seventy *S*s, undergraduate students, tested in a group-situation, were instructed to produce a series of 'heads' (*H*) and 'tails' (*T*) such as they might expect to occur if an unbiased coin were tossed in an unbiased manner for a total of 300 independent tosses. Every *S* responded by filling in appropriate spaces on two IBM answer sheets, each with spaces for 150 responses. 'Head' was indicated by filling in the first space on the answer sheet and 'tail' by filling in the second space. The *S*s were told to work rapidly and continuously until all the spaces were filled.

Results: (1) Run-consistency. The measure of individual consistency used in this analysis was the corrected split-half reliability coefficient based on correlation between number of runs in the first and second halves of the series of 300 responses. The number of runs is a measure representing the frequency of alternations in the series. It is also a function of the length of runs produced, since more long runs result in reduction of the total number of runs. The split-half reliability coefficient for number of runs was found to be 0.89. This indicates that individual *S*s tend to be consistent in the frequency of alternation when attempting to generate a random series.

(2) Consistency of triplets. A second analysis was based on eight different scores representing the frequency of occurrence of *H* or *T* after particular combinations of the two previous responses in the series, *i.e.* the joint probability of a given sequence of three responses. Scores consisting of the frequency of occurrence of each of the possible sequences were obtained for each half of the series. A correlation was computed for each score between the first and second half and corrected by the Spearman-Brown formula to given the reliability-coefficients found in Table I. All of these correlations are significant at the 1% level.

(3) Consistency in number of 'heads.' The number of heads was another measure analyzed for consistency between halves. This analysis was designed to determine whether there is consistency of preference for either *H* or *T* as a response. The corrected reliability-coefficient for number of *H*s was found to be 0.74, significant at the 1% level.

(4) Group-data for runs. The purpose of this analysis was to determine whether there was a tendency for the group as a whole to deviate from chance expectation in their frequency of runs. The number of runs expected in a series of 300 tosses of a coin is 151.[5] The mean number of runs

[4] Ross, *op. cit.*, 136.
[5] H. M. Walker and J. Lev, *Statistical Inference*, 1953, 428-430.

for *Ss* in this experiment was 176. The tendency to produce more runs than expected by chance was characteristic of the majority of *Ss*, as indicated by the fact that 62 of the 70 *Ss* produced more than 151 runs. The deviation from chance was significant at the 1/10% level by a Chi-squared test. The tendency of the group was to avoid long runs and therefore to have more runs than expected in purely random sequences. This finding is in agreement with the hypothesis suggested by Reichenbach, but does not agree with the empirical finding of Ross.[6] This discrepancy will be discussed below.

(5) Triplet-analysis. The eight possible arrangements of triplets were shown in Table I, together with the probability of occurence of each triplet. The eight triplet-arrangements can be classified into three types in terms of alternation and symmetry. The triplets *HHH* and *TTT* contain no alternation between *H* and *T*. Such triplets account for 16% of all the triplets. Of the remaining 84% of triplets in which there is alternation between *H* and *T*, 48% are asymmetrical (*HHT, TTH, THH, HTT*) and 36% are symmetrical (*HTH, THT*).

TABLE I
ANALYSIS OF RESPONSE-TRIPLETS

Triplet	Reliability*	Probability of occurrence†	Alternation	Symmetry
HHH	.73	.08	No	Yes
TTT	.81	.08	No	Yes
HHT	.78	.12	Yes	No
TTH	.78	.12	Yes	No
THH	.75	.12	Yes	No
HTT	.81	.12	Yes	No
HTH	.89	.18	Yes	Yes
THT	.88	.18	Yes	Yes

* Split-half reliability corrected by Spearman-Brown formula.
† Based on frequency of occurrence for the group as a whole.

The group-data suggest two types of tendencies which can be said to characterize the group: (1) avoidance of repetition of the same response; and (2) avoidance of symmetrical response-patterns. These findings are consistent with the findings of other investigators who have analyzed guessing-patterns in other contexts.[7]

[6] Reichenbach, *op. cit.*, 153; Ross, *op. cit.*, 138.
[7] This tendency is also shown of course by the fact that there were more runs than predicted by chance. See L. D. Goodfellow, A psychological interpretation of the results of the Zenith radio experiments in telepathy, *J. exp. Psychol.*, 23, 1938, 601-632; The human element in probability; *J. gen. Psychol.*, 23, 1940, 201-25; B. F. Skinner, The processes involved in the repeated guessing of alternatives, *J. exp. Psychol* 30, 1942, 495-503; *Verbal Behavior*, 1957, 105-106.

(6) H–T analysis. This analysis was carried out to determine whether there is a preference for *H* or *T* for the group as a whole. The mean number of *H*s was 150.5 and the mean number of *T*s was 149.5. These means do not differ significantly from chance expectation. There was a response-preference on the very first trial, however, as about 80% of them were *H*.

Discussion. The main findings of this study are: (1) The *S*s exhibit consistent patterns of responses, judged by coefficients of reliability on a number of measures, when trying to generate a random binary series; and (2) the *S*s, in trying to generate a random series, deviate from randomness by having too many alternations in the series.

There are a number of experimental situations where guessing tendencies, if consistent for individual *S*s, would constitute a source of variance that could be confounded with the effect of other exprimental variables. For example, in experiments on learning where *S* is required to guess which of two lights will go on, the guessing tendencies of the *S* as well as the *E*'s control of reinforcements will determine the response made by *S*. Similar confounding effects may be present in psychophysical experiments and in *ESP* experiments. These consequences have been pointed out before.[8]

The demonstration of individual consistency in guessing-patterns suggests a method that might improve experiments which involve guessing. Given knowledge about response-patterns for a given individual, it should be possible to increase the precision of an experiment by matching *S*s in different experimental groups according to their particular tendencies.

The discrepancy between the present results, showing an excess of alternations, and the findings of Ross calls for an analysis of the two experimental situations. There is one major difference in the experimental method. Ross's *S*s constructed what he believed to be a random binary series by taking cards from either an X-file or an O-file and placing the cards in a center file to represent a random arrangement of Xs and Os. The X- and O-files were placed one at the right and the other at the left side of *S*. Furthermore, it was necessary for *S* to remove each card from the file and stamp it first with an 'X' or an 'O' before he could put it in the center file. The fact that two sides (right and left) and motor operations were involved suggests that *S* might prefer one hand rather than the other, or he might prefer stamping 'X' to stamping 'O.' Any bias in card-selection which favors one side would work against the tendency to alternate. In the present experiment, the use of a preferred hand or preferred side of the

[8] For example, by Goodfellow, *op. cit.*, *J. exp. Psychol.*, 23, 1938, 601-632.

body does not operate. Furthermore the same type of motor response, making a black mark on an answer sheet, is made for both *H* and *T*.

Summary

A set of 300 responses was produced by 70 *Ss*. The set was divided into two sub-sets of 150 responses, and split-half reliability-coefficients were computed for scores of the number of alternations, of particular response-triplets, and of 'Heads.' All of these measures proved to be highly consistent from the first to the second sub-set. For the group as a whole there was a tendency to make more alternations than would be expected in a random series.

THE EFFECT OF DISTRIBUTIONAL SKEWING UPON JUDGMENT WITH FREE CHOICE OF SCALE

By Samuel Fillenbaum, University of North Carolina

When a subject (S) is required to make ratings of a series of stimuli in a situation where no standard is provided, he must make use of some subjective scale of judgment. In the approach employed by Helson, the subjective scale is considered to be completely determined by the level of adaptation.[1] The adaptation-level is taken as some weighted mean function of all the stimuli affecting the organism, each stimulus being judged according to its direction and distance from the neutral region. On Helson's theory, it can be predicted that, with positive skewing of the stimulus-distribution, the adaptation-level, which is a weighted geometric mean of the stimuli, will be pulled down with a resultant upward shift in the rating of all the stimuli. Consistent with this prediction is an incidental finding of Parducci.[2] In a recent experiment on size-estimation, he found that a stimulus in the middle of the distribution was judged significantly larger when it was a member of a positively skewed distribution than when a member of a negatively skewed distribution covering the same range.

The present study, which examined the effect on judgment of varying the distribution of stimuli, is an extension of this work to a situation with less severe distribution-skewing,[3] and to a situation where changes in the judgment of all the series-stimuli may be evaluated. There was one further, important, modification: Ss were permitted free choice regarding the number of categories in the rating-scale. This was done to make for greater generalizability of results, and to minimize possible artifacts resulting from the imposition of a scale with a specified number of categories.[4] Also, this

---

* Received for publication July 14, 1958. Research done under auspices of Defence Research Board of Canada, Project No. D77-94-35-55, at McGill University.

[1] Harry Helson, Adaptation level as a basis for a quantitative theory of frames of reference, *Psychol Rev.*, 55, 1948, 297-313.

[2] Allan Parducci, Direction of shift in the judgment of single stimuli, *J. exp. Psychol.*, 51, 1956, 169-178.

[3] In Parducci's experiment the more frequent stimuli appeared eight times as often as the less frequent ones.

[4] See D. T. Campbell, N. A. Lewis, and W. A. Hunt, Context effects with judgmental language that is absolute, extensive, and extra-experimentally anchored, *J. exp. Psychol.*, 55, 1958, 220-228.

was one way of satisfying the 'equivalence condition' demanded by adaptation-level theory.[5]

Method

The *Ss* (76 men from the Canadian Army who were used in groups of 4-5) judged the slimness-broadness of rectangles projected on a screen singly and in randomized order. Six different rectangles were used as stimulus-objects. Height was kept constant and the ratio of height to base was varied in logarithmically equal steps. The ratio of height to base for the slimmest rectangle was 5.00:1, for the broadest rectangle this ratio was 2.64:1; the change in ratio was always 0.133.

S was allowed to start making judgments whenever he felt ready, and was permitted to use as many categories as he wished in his subjective scale. It was required only that the number zero be used to denote medium and that positive numbers be used for those rectangles considered broader than medium, and negative numbers for the rectangles slimmer than medium.

In all there were 100 exposures, each of about 3-sec. duration, with a second or so between successive exposures. For the first 54 exposures all the rectangles appeared equally often, being randomized in blocks of 6. In the latter part of the session the three slimmer rectangles appeared twice as often as the three broader ones in randomized blocks of nine exposures.

The four judgmental blocks, from Presentation 31 to 54, constituted the data for the first part of the session — rectangular distribution of the stimulus-objects. The five blocks, from Presentation 55 to 99, provided the data for the second part of the session — positively skewed distribution of the stimulus-objects. For every *S* the mean rating given each rectangle during the first and second part of the session was calculated. Every *S* was given a score, the number of rectangles whose judgment shifted up less the number whose judgment shifted down, indicating the net number of rectangles whose judgments shifted in one or the other direction. Also, the value of the adaptation-level in each part of the session was obtained by noting the rectangle nearest which the medium judgment fell. Each *S* was then characterized as either showing no shift, upward shift, or downward shift in adaptation-level.

Results. (1) Judgmental shifts. Over-all, there was no significant shift in judgment upon changing the stimulus-distribution from a rectangular to a positively skewed one. (a) The mean net number of stimulus-objects for which there was an upward shift in judgment was 0.43 (where the maximum possible was 6.00). This yielded a t of 1.13. (b) Of the 76 *Ss*, 34 had net upward shifts in judgment, 27 had net downward shifts, and 15 had no net shifts. The difference between the number showing upward shifts (34) and the number showing downward shifts (27) is not signif-

---

[5] This condition is stated as follows: "The judgment-scale and the stimuli encountered [must be] *equivalent* in the sense that the scale is broad enough to include judgments of all the stimuli encountered and yet is so narrow that its extreme values do not fall outside the range of judgments elicited by any of the stimlui." See W. C. Michels and Harry Helson, A reformulation of the Fechner law in terms of adaptation-level applied to rating scale data, this JOURNAL, 62, 1949, 357.

icant. (c) For each rectangle separately the number of *Ss* showing the various changes in judgment is given in Table I. While judgment of five of the six stimulus-objects does shift upward, for four of them the difference between upward and downward shifts is slight. For Rectangle 2 there is a significant upward shift in judgment ($p < 0.05$ on sign-test) and for Rectangle 6, the broadest rectangle, there is a significant downward shift in judgment ($p < 0.05$ on sign-test).

(2) Shifts in adaptation-level. There was a significant preponderance of downward over upward shifts in adaptation level ($p < 0.05$ on sign-test).

TABLE I

SHIFTS IN THE *Ss'* JUDGMENTS OF 'SLIMNESS-BROADNESS' OF RECTANGLES

Shift	Rectangles					
	1	2	3	4	5	6
Upward	39	47	39	40	37	26
Downward	29	25*	32	32	32	45*
None	8	4	5	4	7	5

* Difference 'up' vs. 'down' is significant as $p < 0.05$

Of the 76 subjects, 40 had downward shifts, 23 shifted upward and 13 showed no changes in adaptation level.

(3) Shifts in judgment and adaptation-level. There was a significant relation between shifts in judgment and shifts in adaptation-level, upward shifts in judgment being associated with downward shifts in adaptation-level, and vice versa. Computed on a 2 × 2 table $X^2{}_{1\ df}$ had a value of 21.8 ($p < 0.01$).

(4) Shifts in judgment, adaptation-level, and number of categories used in the rating-scale. There was no relation between the number of categories used in the rating-scale and either the net shift in the judgment of stimuli, $r = -0.12$, or the shift in adaptation level, $X^2{}_{1\ df} = 0.86$.

(5) Initial location and shifts of adaptation-level. There was no association between initial location of the adaptation-level and any shifts as a function of positive skewing of the stimulus-distribution, $X^2{}_{1\ df} = 1.47$.

Discussion

If the adaptation-level is a weighted, mean function of the stimulus-objects and if judgment is completely determined by the adaptation-level, then, with positive skewing of the stimulus-distribution, one would expect both downward shifts in adaptation-level and upward shifts in the judgment of the serial stimulus-objects. In fact, there was a significant down-

ward shift in adaptation-level without any over-all shift in the rating of the rectangles.

There was no relation between initial location of the adaptation-level and shifts in adaptation-level. Those *S*s whose initial level lay nearest Rectangle 2, the one for which there was greatest upward shift in judgment, were no more prone to show downward shifts in adaptation-level or upward shifts in judgment than anyone else. The significant preponderance of downward over upward shifts in adaptation-level cannot be explained on this basis.

The association between shifts in adaptation-level and judgment is of course consistent with the theory of adaptation-level, but it, of itself, has no necessary implications regarding the determinants of judgment. Consider, if half the *S*s had shown upward shifts in judgment and downward shifts in adaptation-level, and the other half had shown downward shifts in judgment and upward shifts in adaptation-level, then there would have been a perfect inverse relation between shifts in judgment and adaptation-level without the manipulation of the stimulus-objects having any effect upon the average level of judgment or the location of the adaptation-level.

It should be noted that the findings are independent of the number of categories in the judgmental scale. To the extent that our primary concern is with judgment, or changes in judgment of the rectangles, the theory of adaptation-level is inadequate to these data. In this instance the over-all rating of the stimulus-objects remained unchanged even though the neutral region or adaptation-level was sensitive to manipulation of the shape of the stimulus-distribution. Since the same stimulus-series were used throughout and the range was kept constant, our judgmental results are perhaps more consistent with the sort of analysis suggested by Volkmann, who maintains that it is the end stimuli that control the principal properties of the judgmental scale.[6] This approach would, however, have difficulty in accounting for the shift in neutral region.

Summary

In the context of Helson's theory of adaptation-level, the effect on judgment of varying the distribution of the stimulus-objects was examined. *S*s were allowed to use as many categories as they wished in rating the 'slimness-broadness' of each of a series of rectangles. Initially, all the rectangles

[6] John Volkmann, Scales of judgment and their implications for social psychology, in J. H. Rohrer and Muzafer Sherif (eds.), *Social Psychology at the Crossroads,* 1951, 273-294. See also C. W. Eriksen and H. W. Hake, Anchor effects in absolute judgment, *J. exp. Psychol.,* 53, 1957, 132-138.

were presented equally often; later the 'slimmer' rectangles appeared twice as often as the others. Over-all there were no significant shifts in their rating. A significant preponderance of *Ss* showed, however, the predicted downward shift in adaptation-level, and there was also a significant inverse association between shifts in rating and adaptation-level. Insofar as our main interest is the prediction of judgment, or changes in judgment, these results are difficult to interpret by Helson's theory of adaptation-level.

APPARATUS NOTES

AN APPARATUS FOR CALCULATING HISTOGRAMS FROM KYMOGRAPHIC RECORDS

A kymographic record is an historical account of some process, indicating the magnitude of the measured variable at any instant. For some research purposes this is useful or perhaps essential information. There is another account of the variations, on occasion no less useful or essential; namely, a statistical account. The basis of this is a frequency-distribution, graphically a histogram, from which may be calculated statistical indices such as mean, mode, median, and *S.D.* The machine described here (Fig. 1) is designed to extract the data from a kymographic record in a form which lends itself to such statistical analysis. The problem for which the machine was made was to analyze a record of the force used by a rat in a bar-pressing situation.

The principle is simple. Imagine a series of parallel lines drawn above the base line of the record shown in the photograph, so spaced as to represent 10, 20, 30,

FIG. 1

. . . x gm. The length of the x-line under the curve indicates the duration of pressing with a force of at least x gm. The length is greatest for 0 gm. and successively smaller for larger forces. The successive differences indicate the times spent in successive decade ranges of force, and together constitute a frequency-distribution. If rectangles of uniform width and of height proportional to the successive differences are erected on an x-axis of force and a y-axis of time, a histogram results.

In practice it is not necessary to draw lines on the record. The paper is fed on a roller past a transparent cursor calibrated from left to right with lines representing 0, 10, 20, . . . x gm. This cursor can be so moved that any chosen value meets the base line of the record. The reference line (0 gm.) then intercepts the record at the desired height. In the photograph, the cursor is set to measure at 30 gm. The

paper is fed on, to the right in the photograph, until the reference-line touches the record at the point where it passes upwards through 30 gm. The counter, up to now not used, is engaged and is driven as the paper is fed on farther. It is disengaged when the reference-line reaches the point where the record passes downwards through 30 gm. This procedure is repeated throughout the record. The counter sums all the relevant items without the necessity of recording each explicitly. The same method is used to sum the lengths of all parts of the record above any other chosen value.

The roller, with a pressure wheel for securing the paper, is mounted between two endplates which act as guides for the paper. The knob, partly visible in Fig. 1 at the far side, is for turning the roller, thus advancing the paper. The cursor is mounted on a threaded rod and can be moved smoothly and precisely by turning the black disk at the end of the rod. The assembly on the near side, including the counter with its driving wheel on the upper arm, and the actuating key on the lower arm, is held up by the upper spring, keeping the wheel clear of the roller. When the key is pressed the lower spring pulls the counter down and presses the driving wheel on the roller, or the paper when this is in the machine. Turning the knob then drives the counter.

The spring system, rather than simply a key fixed to the upper arm, is used to ensure a steady force on the paper, regardless of the operator's strength of pressing, thus preventing either slip or too great a strain on the counter. It is better to have the friction-wheel running on the paper, rather than adjacent to it on a wider roller. If this is not done and slip occurs because the paper is obstructed, the rotation of the roller will produce a spurious count.

University of Edinburgh J. R. Trotter

THE DECADE-COUNTER TUBE IN THE PSYCHOLOGICAL LABORATORY

The purpose of this Note is to invite attention to the decade-counter tube and to the uses to which it may be put in the psychological laboratory. It may be used for various types of timing as well as for counting.

The principles of operation of three types, decatron, nomatron and remtron, are briefly described in an article by Thomas.[1] In each case a neon type glow is produced on one of 10 cathodes set in a circle around the top of the tube. Unlike most earlier electronic counting devices the conducting circuit changes to a different cathode for each unit of count. This gives the tube great flexibility. A substantial pulse is available at the conducting cathode for operating further circuitry. This was formerly available only in more cumbersome binary counters. Multidecade counting is accomplished by causing the pulse as each 'zero' cathode is struck to operate the next higher decade. Tubes are readily reset to zero.

Reaction-timing in milliseconds will be carried out with little more than the terminal error of less than ±1 m.sec. A mains controlled or crystal controlled 1000-c.p.sec. oscillator is fed into the counter for the period to be measured. The number of cycles counted is the time in milliseconds.

[1] A. B. Thomas, The principles and method of operation of some modern gas-filled counter tubes, *Proc. Inst. Radio Eng., Austral.*, 13, 1952, 311-315.

The number of decades required for timing long periods is not great. Two or three decades will cover fractions of a second in hundredths or milliseconds. Two further decades, one recycling at 6, will time to one minute, a further two to one hour and so on. Power mains control is best used for longer periods.

The use of coincidence-detection enables a pulse to be emitted at any determinable count involving more than one decade. A tube is biased to pass current only when a number of conditions are all fulfilled; in this case, when all decades are on the selected digits. This principle is used industrially for 'batch' counting. At the required count, machinery is put into operation to remove the 'batch' and the counter automatically resets itself to zero. By counting cycles from an oscillator, this principle can be used for the accurate production of a time-interval either with or without resetting. A separate switch is required for setting the count on each decade.

A second and subsequent pulse can be obtained by using extra coincidence detecting circuits. One must be selected to reset the counter if recycling is required. A set of as many switches as there are decades will be required for each output pulse. When the number of decades is large, the number of switches and the panel on which they are mounted soon become unwieldy. It must be pointed out, however, that only two decades are required to allow adjustments of 1% to be effected. The length of the two-decade cycle may be controlled by the oscillator frequency within wide limits and pulse points set as percentages of this period.

A machine utilizing this principle is in operation in the University of Queensland. It will count at speeds in excess of 1000 per second from a pulse of 30 volts regardless of wave-shape. The machine uses four type GS10C decatrons. When four output pulses are required a switch reconnects the tubes as a pair of two decade counters running in parallel. Each two-decade unit provides two output-pulses.

This machine is used in an apparatus designed to examine the limiting conditions for the appearance of the 'phi-phenomenon.' Each light is energized by a bi-stable multivibrator or 'flip flop' circuit. The output-pulses from the counter are fed to the respective grids of the multivibrator tubes. No moving parts are involved. Each unit of count represents five degrees when the counter is set to recycle at the seventy-second count.

The four output-pulses are quite independent and each may be employed to carry out any function. Typical uses include the presentation and adjustment of stimuli and the administration of rewards and punishment. This machine has already given approximately 200 hr. of useful service. No fault or irregularity has developed.

University of Queensland LAURIE E. ENTICKNAP

A SIMPLE AND INEXPENSIVE CARD-CHANGER

The card-changer, described here (Fig. 1), is simple in construction and inexpensive to make.[1] It consists of four parts: (1) a screen, which may be of any desired size, to conceal the experimenter, his materials, and records from *S*; (2) a card-box in which the stimulus-cards are placed in the order they are to be exposed; (3) a device to expose them; and (4) a device to stack them, in order, following exposure.

[1] The apparatus described here was made by Mr. Mayo Grantham, instrument maker in the Psychological Laboratory of The University of Texas.

The dimensions of the exposure-window, the card-box, and the stacking-device are determined by the size of the stimulus-cards selected, which may, within wide limits, be of any size desired. The cards for which the apparatus described here was constructed are 6 cm. high, 11 cm. wide, and 3 mm. thick. They were cut from heavy, white cardboard that they would not bend or crush when exposed in the changer. The exposure-window, centered in the screen at *S*'s eye-height when seated at a table, is slightly smaller, 4.5 × 9.5 cm., that the edges of the cards would be concealed during their exposure. The card-box, in which the stimulus-cards are placed, is slightly larger than the cards, 6.5 cm. high, 11.5 cm. wide (inside dimensions), that the cards would slide freely through it. The box is attached to the screen directly behind the exposure-window. It is 8 cm. deep, that 20 cards of the thickness

FIG. 1. PHOTOGRAPHS OF THE CARD-CHANGER

FRONT

REAR

used could be contained within it. If more cards are desired, the depth of the box may be increased to accommodate them.

The stimulus-cards (C), as shown in Fig. 2, are pressed to the front of the box by a brass plate which is itself pressed forward by two coil springs. The springs are held in place by metal rods which project through metal sleeves at the back of the box and are connected by a bar (D). This bar serves as a handle to retract the brass plate when the stimulus-cards are being put in place.

The exposure-device—shutter and card-changer—is a thin brass plate (B), 11 cm. wide and 22 cm. high, which is suspended from the screen directly above the exposure-window by two spiral springs. This plate runs through grooves in guides (E, E) and extends, when held in a resting position by the springs, from the bottom of the top frame to 0.5 cm. below the exposure-window. A second window (W), 6.5 × 9.5 cm., is cut in the brass plate. This window coincides when the shutter is depressed (*i.e.* the shutter-handle (A) is drawn down as far as it will go) with *S*'s window, thus exposing the card behind it. Two brass dogs (strips R, R), each 1.5

mm. thick, are soldered to the plate at opposite ends of the window (W). When the shutter is depressed, the dogs engage and push the card under them out of the card-box, down into the stacking-device, simultaneously exposing the next-following stimulus-card. The cards, as they are removed, are stacked, as shown in Fig. 1, in the order of their exposure.

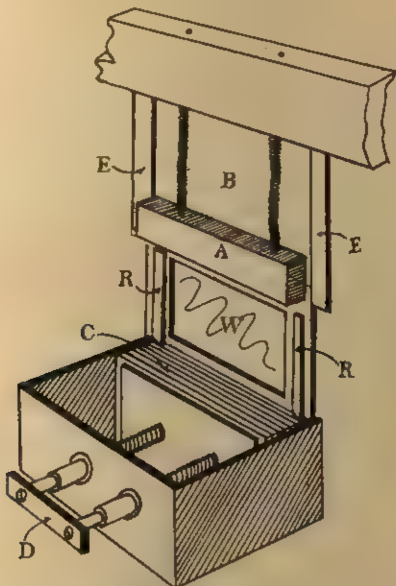


FIG. 2. DIAGRAMS OF CARD-BOX, CARD-CHANGER, AND SHUTTER
A = shutter handle; B = shutter; C = stimulus-cards; D = handle to card-box; E, E = Shutter guides; R, R = dogs to remove stimulus-cards after exposure; W = window in the shutter.

For exposures from 1- to 5-sec. duration, the apparatus may be operated manually with the aid of a metronome or other timing-device. For shorter intervals, a solenoid and an electrical timer should be used.

K. M. D.

NOTES AND DISCUSSIONS

CONVERGENCE, ACCOMMODATION, AND VISUAL ANGLE AS FACTORS IN PERCEPTION OF SIZE AND DISTANCE

Several investigators in the field of apparent size in binocular vision have come to the conclusion that the angle of convergence is a factor in determining the size of stereoscopic images. Early observations on this subject were connected with Meyer's "Wallpaper-phenomenon."[1] More recently, Adams[2] and Herman[3] stated that the geometrical and apparent size of stereoscopic image is related to the convergence of the eyes in the sense that the greater the angle of convergence, the smaller the apparent or estimated size of these images. Accommodation, on the other hand, seems, in their opinion, not to play a role in this respect. It seems, however, that the effect these authors attribute to convergence should be attributed to other factors. For example, they do not take into account the fact that, for fixated objects in normal binocular vision, the very opposite is true; namely, the smaller the angle of convergence, the smaller is the perceived size of the fixated object.

Let us consider the question of relations between convergence and size in the light of other investigations. It has been observed by Helmholtz, Hering, and Zając,[4] that when, in monocular vision as well as in binocular vision with double images, we change the distance of the fixation-point, then other objects placed in the lateral fields of view diminish in size when the fixation-point is moved toward, the observer (*O*) and grow when it is moved away from him. The author's investigations have established the fact that changes in perceived distances from *O* occur at the same time. When the fixation-point comes nearer to *O*, the objects in lateral fields of view moves away from him and vice versa.

From these observations we draw the conclusion that convergence is

[1] H. Meyer, Ueber einige Täuschungen in der Entfernung und Grösse der Gesichtsobjekte, *Arch. Physiol. Heilkunde*, 1, 1842, 316-326. For a description of this phenomenon, see Herman von Helmholtz, *Treatise on Physiological Optics*, trans. of *Handbuch der Physiologischen Optik*, J. P. C. Southhall (ed.), 3, 1925, 316.

[2] O. S. Adams, Stereogram decentration and stereo-base as factors influencing size of stereoscopic pictures, this JOURNAL, 68, 1955, 54-68.

[3] T. B. Hermans, Visual size constancy as a function of convergence, *J. exp. Psychol.*, 21, 1937, 145-161; The relationship of convergence and elevation changes to judgments of size, *ibid.*, 48, 1954, 204-208.

[4] Helmloltz, *op. cit.*, 431; Ewald Hering, *Beitrage zur Physiologie*, 2, 1861, 15 ff.; J. L. Zając, La Localisation en profondeur des images doubles, *Acad. Polonaise des Sciences et des Lettres*, —, 1923, 45-101.

not a necessary factor for perception of sizes of images, and that size is dependent on other factors common both to monocular and binocular vision and varies with the distance of a fixation-point. These factors are accommodation and visual angle. Accommodation determines the size of the retinal image. For positions within the range of accommodation, the nearer the fixation, the greater the curvature of the crystalline lens, the smaller the retinal images, and the smaller also the perceived sizes of the stationary objects in the lateral field of view. This result follows from monocular as well as binocular vision. When, however, binocular and monocular images appear in the same field one must take into account the fact that binocular single images are perceived as increased in size by about 4% in comparison with monocular images. The difference is in horizontal extensions. Kundt's observations of differences in perceived size of images in nasal and in temporal parts of retinas must be taken into consideration here.[5]

When objects in lateral fields of view are situated outside "depth of focus," their images are blurred and larger than corresponding sharp images. Here then is another source of changes in the perceived size of images due to accommodation. Accommodation is, therefore, probably responsible for the phenomena described as 'overconstancy' in various experiments.

Another factor for perception of size of images is visual angle. Two aspects of this factor should be taken into account. (1) When the fixated object is changing its position in space, the visual angle at which it is seen by each eye, also changes. When it moves away from *O*, the visual angle diminishes; when it approaches *O*, the visual angle increases. The consequence of this is that the perceived size in the first case diminishes and increases in the second. The estimated size may change or not according to the law of so-called constancy, but perceived size is really changing.[6] Here also the state of accommodation contributes in the sense described above. (2) The visual angles of objects placed in lateral fields of view do not change with changing distance of the fixation-point, but the geometrical size of the images, measured at the geometrical distance of the fixation-point, grows with growing distance of the fixation. The perceived size of those objects also grows, although in a different way. The effects of accommodation must also be taken into account here. In the purest form, this aspect of the law of visual angle, the so-called "Emmert's Law," can be

[5] August von Kundt, Untersuchungen über das Augenmass und optische Täuschungen, *Poggend. Ann. d. Physik,* 120, 1863, 118-158.

[6] The whole question of size-constancy, which, in my opinion, introduced confusion into the problem of visual depth, and size as it is concerned with estimates of size rather than its perception, is outside the scope of this note.

applied to after-images.[7] Their geometrical size grows with the distance of the fixation-point, not of the background against which they are seen, as Hering contended. This is the correct interpretation of "Emmert's Law."[8] Also the perceived size of the after-image grows with growing distance of the fixation-point, although only within specific limits. Here accommodation has no influence on size, neither geometrical nor percieved. This can be demonstrated by casting an after-image on a background and looking at both, the primary object and its overlapping after-image, through a magnifying glass placed in front of the eye. The size of the after-image will be unchanged, while the object itself would be seen magnified or diminished, upright or inverted, according to the position of the magnifier.

Another factor for the sizes of stereoscopic images, which should be considered, is the relation between the perceived sizes of the two images by the two eyes, which fuse stereoscopically. When these two images are nearly equal in size, the size of the stereoscopic image is dependent on visual angle and accommodation, as described above. When they are unequal, then the relation between the size of these two images is responsible for the size (length) and inclination (slant) to the median plane. The greater the difference between the two perceived sizes, the greater is the geometrical size of the stereoscopic image and also the greater the perceived size. The fused stereoscopic image can become quite large and assume the form of a 'wall.' This relation is the cause of distortions in stereoscopic films and in arrangements with aniseiconic glasses.[9] The same result may be obtained from double images seen in the vicinity of other images or objects.[10]

In case of stereoscopic images, accommodation also plays a role. When one of the two monocular images is sharp and distinct and the other blurred, then the measured size of the stereoscopic image may be much larger than either of the two monocular images, but the perceived size may be the same as the larger of the two. When, for example, the larger image is blurred and the smaller is sharp, then the geometrical size of the stereoscopic image would depend on the relation of sizes of the two images, but the perceived size of this image will be smaller, and no inclination to the median plane would be observed.[11]

---

[7] E. Emmert, Grössenverhältnisse der Nachbilder, *Klin. Monatsbl. Augenhk.*, 19, 1881, 443-450.

[8] Zając, Spatial localization of after-images, *Proc. 15th. Internat. Cong. Psychol.*, 1957, 278-280.

[9] Zając, Depth perception of stereoscopic images resulting from fusion of crossed and uncrossed double images, this JOURNAL, 72, 1959, 163-183.

[10] Zając, Depth perception of double images in the vicinity of other images, *Acta Psychol.*, 12, 1956, 111-129.

[11] Zając, *op. cit.*, this JOURNAL, 72, 1959, 180.

From the above considerations we may conclude that the angle of convergence is not the proper factor on which the perceived size of stereoscopic or double images depends. The influence of accommodation and especially of visual angle in both its aspects is here esssential.

We now pass to the question of depth-perception. With stereoscopic images, our observations were as follows. When, with natural or free stereoscopy and without a material fixation-point, a change of convergence occurs, then a clear relation between the angle of convergence and the perceived distance of stereoscopic images cannot be established. When, however, a fixation-point is introduced differences in distance between this object and the stereoscopic images are immediately perceived. In this case, however, it is not convergence (strictly angle of convergence), since it does not change when a stationary point is fixated, but retinal disparity that is responsible for the perception of depth. In observations with Meyer's wallpapers, we observe very distinctly that the elements of the wallpaper diminish in size as convergence increases, but it is only with great uncertainty that changes in distance are noted. When we place a material fixation-point at the same distance at which the stereoscopic images of the fused elements are seen, we are surprised to discover how near to *O* it must be placed.

It is time to realize, that Berkeley's theories, namely, that convergence and accommodation, as kinesthetic sensations or muscular efforts, contribute to the perception of depth, have no factual basis. The role of convergence must be understood in the sense that, when it is stabilized by fixating some object in the field of vision, the position and sizes of images of other objects are defined, and with them their correspondence or disparity.[12] Changes in fixation bring about corresponding changes in the disposition of retinal images and with them the new geometrical and perceived sizes, distances, and reliefs. Other movements of the eyes also play a role here, not through proprioceptive sensations, but by changes occurring in the size of the images in the two eyes. Experiments carried out by Stevenson Smith showed, that convergence, as proprioceptive muscular sensations, do not have any bearing on the depth of perception.[13] His findings are in complete disagreement with Vernon's, who reported that: "(a) disparity sensations alone can only give rise to a perception of non-equidistance of certain parts; (b) definite impressions of 'nearer' and 'farther,' and the degree of 'nearness' and 'farness,' will be obtained with certainty only when certain other perceptual data can also be employed;

---

[12] *Ibid.*, 179 ff.

[13] Stevenson Smith, The essential stimuli in stereoscopic depth perception, *J. exp. Psychol.*, 38, 1946, 518-521; A further reduction of sensory factors in stereoscopic depth perception, *ibid.*, 39, 1949, 393-394.

kinaesthetic sensations of different degrees of convergence, or of changes in convergence, are probably of great importance also in this respect, particularly over near distances."[14] The facts are quite to the contrary. With binocular vision in the near field of view, the 'nearer,' or 'farther,' or 'equidistantly' are 'immediate, primary data,' related to retinal disparity. There is, moreover, no such thing as 'disparity sensations' corresponding to a perception of non-equidistance. Neither has it been established that, to perceive 'nearer' or 'farther,' one must experience a two-step type of perception mentioned by Vernon. Ittelson's and Ames's conclusions that muscular sensations of convergence and accommodation have something to do with depth-perception in stereoscopic vision are difficult to understand.[15]

Not all the investigators accept the fact that *objective, physical, geometrical,* or *absolute sizes* and *distances* of objects fixated or placed in lateral fields of view *are subject to judgments or estimates, but are not perceived,* and that *relative size and distance, in simultaneous or consecutive stimulations, are subject to perception.*

It is not intended to give, in this short note, a complete analysis of the relation of these different factors with the perceived visual size and distance or relief of objects or images. It is, however, worth mentioning that various attempts have recently been made to explain the perception of size and depth,[16] though none of them has taken all the relevent factors into account.

University of Edinburgh J. L. ZAJĄC

MATCHING TO OBJECTIVE EVENTS IN PROBABILITY-LEARNING—SOME DISCREPANT RESULTS*

In simple probability-learning using a non-contingent series of choices between two alternatives, the usual finding has been that *Ss* approach an

[14] M. D. Vernon, *A Further Study of Visual Perception,* 1950, 100.

[15] W. H. Ittelson and Adelbert Ames, Jr., Accommodation, convergence and their relation to apparent distance, *J. Psychol.,* 30, 1950, 43-62.

[16] Wolfgang Metzger, *Gesetze des Sehens,* 1953, 206 ff., 235 ff., 279 ff., 339 ff.; L. H. Hardy, Gertrude Rand, M. C. Rittler, A. A. Blank, Paul Boeder, *The Geometry of Binocular Space Perception,* 1953, 1-67; J. J. Gibson, *The Perception of the Visual World,* 1950, 59 ff., 92 ff., 100 ff., 137, 163 ff., 175 ff., 190 ff.; A. S. Gilinsky, Perceived size and distance in visual space, *Psychol. Rev.,* 58, 1951, 460-482; Norbert Günther, Die Grundgesetze der Raumwahrnehmung, *Optik,* 2, 1947, 58-80.

* This study was done under auspices of the Defence Research Board of Canada, Project No. D77-94-35-55, at McGill University.

asymptote of responding to an alternative at about the rate that it actually occurs.[1] Estes has interpreted this matching of behavior to events as an example of his statistical contiguity-theory, and other writers have attempted to derive these results from an analysis using game-theory.[2]

The data reported below were obtained from a group of *Ss* different from the usual college sample, and the results are rather different from the common findings.

The situation and procedure we used were essentially the same as those used by Jarvik.[3] The *Ss*, 85 Canadian soldiers who served in small groups of 4 or 5, were asked to predict which of two mutually exclusive alternates would occur on each of a set of 120 trials. They were provided with mechanics' paper and were given the following instructions:

I am going to read you a list of words, it will consist of the words "plus" and "check." Before each word I'll give you the signal "now." With each "now" I want you to guess whether the next word is going to be "plus" or "check." We'll start with a "now." In response you should write either a plus sign (+) or a check (√) in the left top square of your sheet, depending on which you expect. After you have done this I'll give you the right answer and you are always to write this in the square to the right of the mark representing your guess. With each "now" take the next lower square for your guess. Remember the object of this is to guess each sign in advance as accurately as you can, so as to get as many right as possible.

The whole thing will be on a record, the signals "now" will come fairly rapidly. If you miss one, draw a line through the two squares indicating that you missed, and go on to the next word.

For the first 60 trials the ratio of "plus" to "check" was 2:1, for the last 60 trials this ratio was changed to 1:2. Trials were randomized within blocks of nine, and were presented at intervals of 5 sec.

Results. The data were analysed in blocks of 10 trials. Table I gives the mean number of plus-responses in Blocks 2–12; the first 10 trials were not scored as a number of *Ss* initially had difficulty getting into step with the record. Reference to the bottom row of Table I shows that, for the group as a whole, the rate of responding to the preponderant alternative does increase over the first six blocks. The rate on Trials 51-60 is, however, still short of the rate of occurrence of this alternative. Decrease in probability of a plus during Trials 61-120 leads initially to a slight increase in rate of responding but over the last three blocks there is a decline in rate of response to a level a little below 50%. The terminal rate is still considerably greater

---

[1] See, for example, D. A. Grant, H. W. Hake and J. P. Hornseth, Acquisition and extinction of verbal expectations in a situation analogous to conditioning, *J. exp. Psychol.*, 42, 1951, 1-5, and H. W. Hake and Ray Hyman, Perception of the statistical structure of a random series of binary symbols, *ibid.*, 45, 1953, 64-74.

[2] W. K. Estes and J. H. Straughan, Analysis of a verbal conditioning situation in terms of statistical learning theory, *ibid.*, 47, 1954, 225-234; H. A. Simon, A comparison of game theory and learning theory, *Psychometrika*, 21, 1956, 267-272.

[3] M. E. Jarvik, Probability learning and a negative recency effect in the serial anticipation of alternative symbols, *J. exp. Psychol.*, 41, 1951, 291-297.

than the probability of the event. In neither instance is there a very close approach to matching of responses to the objective events.

For further analysis, the *Ss* were divided into two groups. The first, Group A comprised those making more than 10 plus-responses on Trials 41-60, and the second group (Group B) comprised the *Ss* making 10 or fewer plus-responses during this period.

It can be seen from Table I that during Blocks 2–6 the *Ss* in Group A showed a considerable increase in rate of responding to the preponderant alternative and then, with shift in the probability of this event, a considerable decrease in rate of

TABLE I

MEAN NUMBER OF PLUS-RESPONSES PER BLOCK OF 10 TRIALS

Group	Block					Block					
	2	3	4	5	6	7	8	9	10	11	12
A (N=43)	3.86	5.05	4.72	6.05	6.84	6.37	5.84	5.84	5.21	5.35	4.86
B (N=42)	4.17	3.97	4.21	3.93	4.47	5.38	5.52	6.00	5.16	5.59	4.78
Total (N=85)	4.01	4.51	4.46	5.00	5.70	5.89	5.69	5.92	5.19	5.48	4.82

response. The *Ss* in Group B, on the other hand, showed little shift in rate during Blocks 2–6 and then, with a decrease in probability, showed an initial rise in rate of responding followed by a decline. For this group, at the end, the rate of responding with a probability of a plus of 0.33 was still a little higher than the rate of responding after 60 trials with the probability of 0.67. It is as though *Ss* in the one group were sensitive to the probabilities with which events occurred and to shifts in these, while *Ss* in the other group were not.

The higher the rate of responding on Blocks 5 and 6 the more likely the *S* is to show a decrease in rate of response with shift in probability. Thus, the coefficient of correlation between the number of plus-responses on Blocks 5 and 6 and the decrease in this number from Blocks 5 and 6 to Blocks 11 and 12 is 0.73 ($p < 0.01$). Of the 43 *Ss* who made more than 10 plus-responses during Blocks 5 and 6 only five made more such responses on Blocks 11 and 12, while 35 made fewer, and 3 made the same number. On the other hand, considering the 42 *Ss* who made 10 or fewer plus-responses on Blocks 5 and 6, we find that 30 increased the number of plus-responses on Blocks 11 and 12 (when there had been a decrease in objective probability), 10 showed no change, and 2 gave fewer such responses. A chi-square computed from these data is highly significant ($p < 0.01$) although this measure is of doubtful usefulness in this situation.

There was no correlation between rate of responding on Blocks 5 and 6, or shifts in rate of responding with change in objective probability, and intelligence as measured by over-all score, or vocabulary or analogy subscore, on the Canadian Army M test.

Discussion. Regarding the group as a whole there is some adjustment of rate of response to the probabilities of objective events and shifts in these, but the discrepancy from perfect matching is considerable. For example, on Trials 41-60 of a 0.67 schedule, 30 of the 85 *Ss* are still responding at a rate less than 50%. Dividing the *Ss* into those responding at better than 50% on Trials 41-60, and those not, it is seen that the *Ss* in the former group are more sensitive both to the objective probabilities and shifts in

them. In a sense, they respond more appropriately.

There is one important difference between this study and the various studies that have found almost perfect matching. These used college students as *S*s (generally volunteers), whereas in this study the *S*s were army personnel with less formal schooling and probably lower average intelligence, who had been sent to the laboratory for a half day's testing.

Some experiments by Edwards involving money rewards, also on college students, yielded data counter to the generalization that responses match objective events.[4] In his case, however, the asymptotic rate of responding was always more extreme than the probability of the event. To account for this Edwards resorts to an analysis deriving from decision-theory. Such an analysis can hardly handle the data of this experiment where the final rate of responding in both instances was less extreme.

Without much conviction two quite opposite possibilties might be raised. It may be that, for our *S*s, there is no systematic effect of the probability of the objective events. The partitioning of *S*s into two groups may have capitalized on chance, and the downward shifts of those responding above 50% and upward shifts of the others might just be regression effcets. On the other hand, it may be that there are systematic effects on the response of all the *S*s. Those in Group B might just learn more slowly; the paradoxical increase in rate of responding of these *S*s upon decrease in probability of an event might be compatible with such an interpretation.

The data reported here which are somewhat at variance with the common findings suggest the possible importance of subject and individual difference variables, and indicate that some qualification may be necessary in those analyses which predict matching of events in the simple situation of probability learning.

University of North Carolina — SAMUEL FILLENBAUM

EYE-DOMINANCE AND HEAD-TILT

While testing several individuals for eye-dominance, the author observed that the heads of left-eyed subjects were tilted slightly to the right and, contrariwise, that the heads of right-eyed *S*s were tilted slightly to the left. This postural imbalance was especially noticeable when the *S*s stood erect during the test of eyedness. To check these observations on a larger sample, the eyedness and posture (head-tilt) of 163 students from Introductory Psychology were examined. Group-tests were used for both eyedness and head-tilt. These tests were administered in a large lecture hall.

[4] Ward Edwards, Reward probability, amount, and information as determiners of sequential two-alternative decisions, *J. exp. Psychol.*, 52, 1956, 177-188

For the test for eyedness, a new group-test was used.[1] *S* was instructed to sit erect and to fixate a point placed upon the blackboard before him. While fixating this point, he was instructed to bring a pencil, which until then had been held vertically with his right [left] hand at his nose into line with the fixation-point. He was then told to close his right [left] eye and to note whether the pencil was still in line with the fixation-point or had shifted to the right or to the left of it. Eight trials, four with the pencil held in each hand, and four with each eye being closed, were required of every *S*. If the pencil remained in line when the right eye was closed, *S* was judged to be left-eyed; if it jumped to the right, he was judged as right-eyed. If it remained in line when the left eye was closed, he was judged as right-eyed; if it jumped to the left, he was judged right-eyed. If fewer than six of the eight trials were inconsistent with these results, *S* was judged "ambiguous."[2]

For the test of head-tilt, *S* was instructed to sit as straight as he could and again fixate the mark on the blackboard. He was then told to hold the pencil vertically about 10 in. in front of his nose while continuing to fixate the point on the blackboard—a position in which he would see two images of the pencil. He was then asked to note which of the two images, if either, appeared higher. Two trials, with the pencil being held in each hand, were required from every *S*. If the left image was higher on both trials, then a right head-tilt was scored. If the left image was higher on one trial but both images were level on the other trial, then a tendency toward a right tilt was scored. If the right image was higher on one trial but both images were level on the other trial, then left tendency was scored. If the images were level on both trials, or the opposite image was higher on the second trial, or there was inability to see double images, then 'ambiguous' was scored.[3]

The accompanying contingency table summarizes our results.

	Eye-dominance		
	right	left	ambiguous
left	31	2	4
left tendency	10	3	3
right	14	21	10
right tendency	10	2	4
ambiguous	20	10	19

[1] Devised by Mr. D. Friedheim and Professor K. E. Zener.

[2] My thanks are due Mr. Herbert Crovitz and Mr. William Groman for assisting in this testing.

[3] The author is indebted to Professor Zener for suggesting this test of head-tilt.

The table indicates a definite correlation between eyedness and head-tilt as they are measured in this study. Chi-square for independence is 36.2, significant beyond the 1/10% level of confidence for $df = 8$. Thus a right-eyed *S* tends to carry his head to the left and a left-eyed *S* tends to carry his head to the right, and *Ss* with ambiguous eye-dominance tend not to manifest any consistent head-tilt. The clarity of this relationship is obscured somewhat by an over-all cultural tendency to tilt the head to the right and it is, therefore, possible that a quantitative estimate of amount of head-tilt would provide an even better discrimination between the eyedness-groups.

One 'functional' interpretation of these results is that the eye used for alignment in the eyedness-test is often unconsciously carried closer to the midline of the body—resulting in a more accurate alignment of the entire individual. No causal explanation for the origin of either the postural or visual imbalance is presently attempted. Since, however, it is probable that specification of eyedness as well as handedness will become more imperative in experiments involving lateral anistopropies in visual perception, and since selection of eyedness-groups on the basis of individual tests can be laborious, it is thought that the demonstration of a group-test which rapidly yields reliable results that correlate with other personal characteristics would constitute a timely contribution.

Duke University GEORGE GREENBERG

ATTNEAVE'S INTEROCULAR COLOR-EFFECT

In a recent number of this JOURNAL, Attneave reported an observation made with after-images from which he concluded that there must be occurring some central binocular interaction.[1] The subject of binocular interactions has always been a contentious one and many claims for their existence have been shown to rest on poor evidence. Such, as I believe, is the case with Attneave's claim.

Attneave produced an after-image of a saturated color in one eye and then, quickly opening and closing the eyes in alternation, he saw after-images alternating in color between the original color and its complementary. He concluded, therefore, that some central, binocular interaction must be occurring. Several facts are relevant in correcting this conclusion.

First, as is well known and as Summer and Watts attest,[2] an after-image in one eye may be seen when the other eye only is open and it is usually

[1] Fred Attneave, An interocular color-effect, this JOURNAL, 70, 1957, 318-319.

[2] F. C. Sumner and F. P. Watts, Rivalry between monocular negative after-images and the vision of the other eye, this JOURNAL, 48, 1936, 109-116.

the same color as the stimulus with a darkish halo round it. This result is not actually dependent on the open eye but is merely observed as if in the space seen by the open eye. Some of the dark field of the closed eye is also similarly 'projected.'

Secondly, as Robertson and Fry show,[3] an after-image produced in both eyes, or one eye, when projected onto a dark surface, is seen as the same color as the original stimulus; when projected onto a white surface it is of the complementary color. When both eyes are closed it is as if there were a dark surface present and the after-image is again the same color as the stimulus.

Thirdly, as is also well known and as Asher shows,[4] a flickering light enhances the clarity and duration of after-images. When Attneave alternately opens and closes the eyes, he is in effect producing a flicker and the after-images are seen in striking fashion. When the eye with the image is open, the complementary image is seen; when the other eye is open an image of the same color as the stimulus is seen, for the reasons given above.

To test my explanation of Attneave's effect and to show that binocular interaction is not necessarily involved, three simple experiments will suffice.

Experiment I. A colored after-image is produced in one eye, then that same eye is opened and closed while the other eye remains either opened or closed. Attneave's effect is still seen. It has nothing to do with what the other eye is doing.

Experiment II. A black sheet of paper is juxtaposed to a white sheet. A colored after-image is produced in one eye. That same eye is opened and the other kept closed. If the gaze is alternated from the white to the black paper, Attneave's effect is seen again. It has nothing to do with opening and closing the eye. It is the alternation of dark and white ground upon which the image is projected which is important.

Experiment III. A colored after-image is produced in one eye and a black sheet of paper held before the same eye. If the eyes are alternately opened and colsed as in Attneave's experiment, *no* alternation of after-images will be seen, the image will remain the same color as the stimulus.

Thus the alternation of the opening of the eyes is neither a necessary (Experiments I and II) nor sufficient condition (Experiment III) for Attneave's reported effect, which can not therefore depend on binocular interaction.

University of Durham, England IAN PORTEOUS HOWARD

[3] V. M. Robertson and G. A. Fry, After-images observed in complete darkness, this JOURNAL, 46, 1937, 265-276.

[4] H. Asher, Contrast in eye and brain, *Brit. J. Psychol.*, 40, 1950, 187-194 (see Experiment 8).

SIXTY-SEVENTH ANNUAL MEETING OF THE AMERICAN PSYCHOLOGICAL ASSOCIATION

The 1959 meeting of the American Psychological Association was held in Cincinnati, Ohio, on September 3-9. There were about 5900 persons in attendance. In accordance with preferences indicated by a survey of the membership of the Association, emphasis at this meeting was placed on symposia, invited addresses, round tables, workshops, etc., as opposed to the presentation of individual papers. There were consequently more of the broad type of events and fewer papers on the program than in recent years.

'APA Day,' now well established on the program, began with addresses by Drs. Frank A. Beach, Paul E. Meehl, and B. F. Skinner, recipients in 1958 of the Association's Distinguished Scientific Awards. Then the recipients of the 1959 awards, who will give their addresses at the 1960 Convention, were announced. They were: Dr. Leon Festinger, Stanford University; Dr. Donald B. Lindsley, University of California at Los Angeles; and Dr. Neal E. Miller, Yale University. The American Psychological Foundation's Gold Medal was awarded and received by Dr. E. G. Boring of Harvard University, and the Foundation's Science Writer's Prize was awarded to Miss Marjorie van de Water of Science Service. In the afternoon, Dr. Robert S. Morison, Director of the Division of Medical and Natural Sciences, Rockefeller Foundation, gave an address on "Gradualness, gradualness, gradualness—I. P. Pavlov." The 'Day' was closed by Dr. Wolfgang Köhler's presidential address on "Gestalt psychology today."

At the annual business meeting, Dr. Donald O. Hebb of McGill University was installed as president of the Association for the forthcoming year and the announcement was made that Dr. Neal E. Miller of Yale University had been chosen as president-elect. It was also announced that Dr. Roger W. Russell had resigned from the executive secretaryship of the Association and that Dr. John G. Darley, University of Minnesota, had been elected as his successor. Other officers for 1959-1960 are: Dr. Wolfgang Köhler, past president; Dr. Launor F. Carter, recording secretary; and Dr. Meredith Crawford, treasurer. Newly elected members of the Board of Directors for 1959-1962 are: Drs. Victor Raimy and Quinn McNemar.

For many years past, Convention Managers have been appointed by the Board far in advance of the meetings that the myriad details may be arranged. In recognition of the tremendous task involved and the need for continuity, a permanent post, Convention Manager, was created. Professor

George S. Speer was elected to this office and Dr. Philip Ash was apopinted Assistant Manager.

Dr. Howard B. Lyman, Convention Manager of the Cincinnati meetings, and the members of the Local Arrangements Committee spent many hours in planning the 1959 meetings. Their efforts were rewarding as the Convention was well organized and smooth-running. Registration proceeded apace; there were at no time long waiting lines, even during peak periods. The exhibits were many and varied and the exhibitors were well pleased with the arrangements. The facilities provided the Placement Office accommodated the ever-growing number of employers and applicants. A dance, held Saturday evening, September 5, was a great success with members, young and old, dancing to conntinuous music provided by several orchestras. The child-care facilities were excellent. It is hoped that similar arrangements will be made at future conventions.

A new feature this year was a series of conversational hours during which younger members were given opportunity to meet and to talk with their older and better known confrères. These sessions proved popular as they provided opportunity for discussion in an atmosphere less formal than programmed events.

The 1960 Convention will be held in Chicago, Illinois, September 1-7, 1960.

American Psychological Association
Washington, D.C.

JANICE P. FISH

ERRATA

Dr. J. L. Zając invites attention to the following errata that occurred in his study on "Depth perception of stereoscopic images resulting from fusion of crossed and uncrossed double images," which appeared in the June, 1959, number of this JOURNAL (Vol. 72, pp. 163-183).

P. 173, the first word in the fifth line from the bottom of this page should be "image" instead of "rod."

P. 174, Formula [4] should be:

$$Y_{F_1} = \frac{2Y_{A_1}Y_{B_1}}{Y_{A_1} + Y_{B_1}}, \text{and}$$

Formula [5] should be:

$$Y_{F_2} = \frac{2Y_{A_2}Y_{B_2}}{Y_{A_2} + Y_{B_1}}.$$

P. 175, second line, the plus (+) and minus (—) signs should be reversed to read "in the first use '—' is used and in the second '+'."

Formula [9] should be:

$$X_1 = \frac{d(Y_{B_1} - Y_{A_1})}{Y_{A_1} + Y_{B_1}},$$

and the designation of Formula [10] should be bracketed.

P. 177, Footnote 8, the reference is *Psychol. Forsch.*

P. 181, ninth line down from the top, "also limited between" should read "also limits between."

P. 182, the first sentence of the Summary should read "Depth perception and relief, resulting from stereoscopic fusion of crossed and uncrossed double images."

K. M. D.

An Acknowledgment

The portrait and signature of Professor Agostino Gemelli appearing on the frontispiece of this number were received through the courtesy of Professor Leonardo Ancona, a former student and colleague, whose necrology of Professor Gemelli appears upon the following pages. The photograph, which was autographed, was taken in 1957 when Professor Gemelli was seventy-nine years old.

K. M. D.

Agostino [Edoardo] Gemelli: 1878-1959

On July 15, 1959, aged 81 years, 6 months, and 3 days, Father Agostino Gemelli, who had been ill since the fall of 1958, died in Milan, Italy. His death is a grievous loss to Italian and European psychology. American psychologists, who have had occasion to coöperate with Dr. Gemelli, will also feel the bereavement. He was born in Milan on January 18, 1878, into an old Lombardy family of comfortable means and of liberal views. He was christened Edoardo.

After a secular, classical education in Milan, he entered the faculty of medicine at the University of Pavia in 1898. There he followed with particular interest the work of the neurophysiologist Camillo Golgi, and later became his assistant. He acquired from Golgi not only the fundamentals of the experimental method but also great finesse in the manipulation of apparatus. His skill with instruments, despite his very large hands, was so great that his delicacy and nimbleness often aroused the astonished admiration of his students.

During his study of medicine, he became involved (as a vigorous youth of action as well as of thought frequently will) in the political and social issues in ferment of his time. He took an active part, and founded and directed Marxist periodicals and wrote polemical articles for the newspapers of that day, despite his liberal leaning. He discharged his military obligations, after receiving his M.D. degree in 1902, as a physician at the Sant'Ambrogio Hospital in Milan. He served his tour of duty there with his old friend Ludovico Necchi. This close relationship with Necchi had a profound spiritual influence upon him. It probably played an important part in his sudden entry, in 1903, into the Franciscan convent of Rezzato. He became a full member of the Order in 1907 and an ordained priest the following year when he changed his first name from Edoardo, which he dropped, to Agostino.

From then on, Padre Agostino put his intellectual and scientific efforts at the service of his Church. In 1909 he founded the *Rivista di Filosofia Neoscolastica* in opposition to the periodical *Critica* headed by Benetto Croce. In an effort to encourage a return of the cultural to a theocentric position, he also assisted in the founding of the journal *Vita e Pensiero*.

Following his specialization in the histology of the nervous system under Golgi—which disappointed him as it did not lead to the understanding of human life as he had expected—Gemelli turned to research of a surgical type on the encephalon of the higher animals. He investigated the phenomenology of the emotional processes in an attempt to find their

locus. It was a short step from there to the investigation of the psychical functions.

In this evolution, he was greatly helped by his religious conversion which brought him from a position of active anticlericalism of an openly Marxist kind to a full acceptance of the spiritualistic tenets of Christianity. His conversion led him to the belief that man should be considered as a whole with a thorough integration of his psychical functions or activities. This point of view, which he held throughout his long life, explains the different stages in his scientific development.

He obtained his Ph.D. degree in psychiatry from the University of Pavia in 1904 and then, following his training and ordination as a Franciscan priest in 1908, he went for three years (1911-1914) to the University of Munich to study with Kraepelin and Külpe, both outstanding students of Wundt, to learn the principles and methods of introspection, which were then emanating from Külpe's School of Imageless Thought; in fact, they seemed to Gemelli to be superior to Wundt's "brass instrument" psychology in which he had previously been steeped and had supported and spread.

After his habitation at Munich as a *Privatdozent,* he went (1914) to the University of Turin as a *Lettore di Piscologia* to work with Kiesow, another outstanding student of Wundt. At this time he was invited to the chair in psychology at the University of Tokyo in Japan, but, due to the First World War in which he soon found himself involved as a chaplain, physician, and psychologist, he declined. He remained at Turin until 1922, when he founded the *Università Cattolica del Sacro Cuore* in Milan, of which he became the *Rettore Magnifico* (president) and professor of psychology. He retired from formal teaching in 1953 but continued as *Rettore* and with his writing and experimental work in psychology until his last illness.

Throughout his professional life, Gemelli followed closely the results of psychological researchers in America—and for many years he was the only Italian to do so. He not only reported these studies in Italian periodicals but also initiated researches to check and to elaborate the results brought to light by his colleagues from across the Atlantic. Gemelli must also be credited with the survival of psychology in Italy, as there was a time, during the Fascist regime, when anyone who called himself a psychologist and tried to carry out experimental work in that field was looked upon with utmost suspicion.

Gemelli accepted Lersch's anthropological doctrine, soon after its pronouncement, as it seemed to him to afford the most satisfactory point of view for an all-embracing concept of man. From Lersch's insight, Gemelli

proceeded to formulate his own concept of psychology—the 'personalistic point of view'—and the establishment of his own school of thought.

According to Gemelli's school of thought, man is, first of all but not entirely, a biological datum. All the methods and conceptions designed on a scientific basis and yet purporting to manage a study of his activity, must therefore take into account this fundamental biological dimension or, at least, they must not clash with it. In addition to this, there is another fundamental datum of human life; namely, its 'interiority' which is strictly personal. This latter aspect is not to be conceived as a mere function of the biological structure for it must be recognized as being independent and innate. Unless the personal dimension is considered, the study of human activity cannot be regarded as complete, just as, without reference to biological aspects, psychological research cannot be said to be scientific.

Gemelli combined these two aspects of man in his concept of 'personalistic orientation' (*soggettività*), which he placed at the center of his concept of psychical life. Personal is the body but no less and perhaps even more so is the *ego,* which Allport calls the "propriate functions." Both body and ego must, as Gemelli held, be regarded as in an intimate interaction of a functional nature from which there emerges, full and complete, the specific quality of human activity. Gemelli centered his scientific activities and those of his students around this 'personalistic orientation' (*soggettività*), as the work from his laboratory clearly shows.

In respect to the perceptive function, he took the precise measurements of physiological optics, of optical-geometrical illusions, of electroretinography, and of electroencephalography as his starting point and inserted them into the frame of reference of intentionality, meaning, and personal importance that such functions have. For the investigations of motivation, the body schema, and orientation in space, he brought the objective data back again to the main interest, *i.e.* to the existential purpose that they had for the subjects that manifested them. This procedure led him beyond the mere objective data to an understanding of more general and meaningful laws and functions.

In respect to language, Gemelli conducted researches in electroacoustics that are still regarded as models of electronic techniques. He did not, however, stop to investigate the play of single phonemes and words in their different succession but viewed the play in function between the contingent mental state of the subject and his unique individuality. Because of this method, the phenomena of social life have been interpreted by Gemelli not only as a function of a natural dynamic 'field' but also as the dynamism of subjectivity influenced, but only in part determined, by the dynamics external to the personal limits.

Lastly we note that, in the field of his researches in criminology, Gemelli rejected every biological or socially deterministic concept to prove the need and the usefulness of such a study of behavior as may be able to take into account the drive-character of psychical activity.

With knowledge of Gemelli's personalistic point of view, it is easy to understand why he never accepted the principles of pure behaviorism—neither the biological behaviorism advocated by Watson nor the sociological espoused by Lewin. He thought that both of these schools tended to make man the object and not the subject of the motivational forces. Nor could he accept the tenets of *Gestaltpsychologie,* which he regarded as an interpretation of partial facts that are too intellectualistic and formal to be used to interpret the higher facts of psychical life and the interpersonal reactions related to them.

It is also easy to understand why Gemelli upheld the independence of psychology from biology, on the one hand, and from sociology on the other hand.

It can, furthermore, be now understood why Gemelli offered such resistance to Freud's psychoanalytic view as a global interpretation of psychical life in which personality is reduced to a secondary manifestation of the libido. For reasons strictly connected with his own personality, Gemelli feared the psychoanalytical interpretation of human life; yet he recognized that psychoanalysis contributed fundamentally to the field of affectivity.

The range of activities in applied psychology covered by Gemelli and his students was very wide. It extended from psychotechnics to vocational guidance and from industrial counseling to clinical psychology. He was the first in Italy to promote vocational selection of soldiers and airmen, establishing the procedures at the Supreme Headquarters of the Italian Army during the War of 1914-1918.

The list of his writings is imposing and it contains several hundred items. In addition to his writings, he established in 1920 in coöperation with Kiesow and an illustrious Board of Associate Editors (Benussi of Padua, Botti and Ponzo of Turin, Colucci of Naples, De Sanctis of Rome, and Morselli of Genoa) the *Archivio Italiano di Psicologia;* and in 1940, he founded, and edited until his last days, the *Archivio di Psicologia, Neurologia e Psichiatria.* The Institute of Psychology, which he organized at his University in Milan, is the best appointed and equipped in Italy. A Postgraduate School of Psychology, organized along the line of similar schools in the United States, is attached to the Institute.

Leonardo Ancona

Catholic University of Milan

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Psychotropic Drugs. Ed. by S. GARATTINI and V. GHETTI. Amsterdam, Elsevier Publishing Co. and Princeton, D. Van Nostrand Co., 1957. Pp. 606. $19.50.

Psychopharmocology Frontiers. Ed. by NATHAN S. KLINE. Boston and Toronto, Little, Brown and Co., 1959. Pp. xxxii, 533. $10.00.

The Effect of Pharmacologic Agents on the Nervous System. Research Publications Assoc. for Research in Nervous and Mental Disease. Vol. 37, Ed. by FRANCIS J. BRACELAND. Baltimore, Williams and Wilkins Co., 1959, Pp. 488. $13.50.

These three large volumes, representing contributions from nearly 300 scientists and clinicians, are reports of papers read and discussions transcribed at three major 1957 meetings: the International Symposium on Psychotropic Drugs at Milan in May; the Psychopharmacology Symposium of the Second International Congress of Psychiatry in Zurich in September; and the proceedings of the ARNMD meeting in New York in December. The aim of all of them being the dissemination of information sufficiently precise to allow some tentative generalizations about the effect of the new drugs on human behavior.

To assess the present state of knowledge of brain function and of the modification of behavior by drug action is a difficult task—so difficult that the Royal Medico-Psychological Association in its 1959 *Recent Progress in Psychiatry* found the breakthrough so promising, but the results so confusing, that the entire subject was omitted as too complex to allow ready summary.

In the present volumes, the biochemists, the pharmacologists, and the psychiatrists have gathered in the three countries to read basic papers, share experiences, and check progress. The psychologists are also represented, with contributions of laboratory techniques basic to the assessment of behavioral effects and of research designs allowing of possibly valid deductions. The first two volumes are largely devoted to trying to find an answer to the worth of the tranquilizers. Despite the ubiquity of the witticisms about using the drug of the week before its potency is lost, the new drugs have manifest and important influences on behavior, and most of these influences still need objective assessment. The broader problems involved in the whole range of inter-relationships between drugs and behavior are the concern of the third volume.

The drugs are presently classified as the neuroleptics, such as chlorpromazine and phenothiazine derivatives; the sedatives and tranquilizers, such as meprobomate; the hypnotics, such as phenobarbital; the psychotonics, such as amphetamine or caffeine; the euphoriants, such as alcohol, opium, or cocaine; the hallucinogens, such as mescaline; and the depersonalizers, like lysergic acid diethylamide. Each of these classes typifies certain clinical actions and phamacological features, involving possible differential action on central or autonomic nervous system.

Psychologists who have been highly concerned that we have not yet developed sufficiently objective measures of the effects of psychotherapy, among other things,

will be surprised to learn that the internists and pharmacologists have been able to devise *no* objective test for measuring the effectiveness of therapy for even the anticonvulsants, because the same global phenomena may result from quite different mechanisms. The criteria for drug action, while biochemical and physiological in the initial tests, are behavioral in the last analysis.

Techniques for the assessment of behavioral change include the objectification of clinical scrutiny by rating scales, but there are other less global, more exact, but hardly less inadequate measures. Physiological assessment should include animal assay and generally electroencephalographic assay. The neurological signs, which are often behaviorally interpreted as "side effects," can also be regarded as essential prerequisites of behavioral change—the parkinsonisms introduced by some drugs may well be a clinical sign that the maximum useful therapeutic level of a drug has been reached.

The area where psychologists should make a most significant contribution is in the study of the placebo response, which is now recognized as a much more complex reaction than just a mode of control in a pharmacological experiment. What range of social, ethical, physiological behaviors are relevant to such responsivity is quite unknown. And yet Stewart Wolf points out, "Placebo effects are probably the most relied upon aspects of pharmacotherapy today, however unintentional this may be on the part of the physician." It is not that the behavior of the human subject is unpredictable when administered a drug. It is that there is a variability in effect following drug ingestion presumably "due to opposing or reinforcing forces growing out of the meaningful situations from day to day in the life of the subject." For hypnotic and sedative drugs, the result in the words of Louis Lasagna is, "Despite the greatly increased use of more precise and imaginative neurophysiologic techniques, despite the considerable amount of biochemical work on these compounds, despite the tremendous enthusiasm for experimental psychologic methods in the study of sedatives and hypnotics, I do not believe that we have any but the foggiest notion of how these drugs act."

So the criterion problem that makes validation of so many global psychological assessment procedures so difficult is really no different in medicine. The experimental method leads to the clinical trial, which, over a variety of situations and over some time, is the payoff in estimating effects of drug action. This results from the complexity of problems for which no one has been able to devise either experimental control or statistical interpretation subtle enough to give answers. But it is also true because the operational criteria in psychiatry are ill-defined and sometimes less than tidy. David Rioch points out that as the experimental pharmacological methods grow more precise and more limited, the major need in the clinical, social, and behavioral disciplines is for "a conceptual frame of reference and a terminology which will adequately represent the phenomena and the operational manipulation of all the phenomena representing research in all these fields." As long as investigators operate on a "best judgment" principle or on the basis of some consensus, the concepts in use will represent "socially preferred value systems, models and metaphors, though often couched in terms which provide a klang-association with the terminology of the physical sciences." So "the limitations in the social and behavioral sciences are those imposed by the lack of operationally definable units of behavior and units of communication . . . as in the better systematized sciences, the unit needed for study of a problem is a function of the experimental objective and it

needs to be selected on economic principles—namely, on the basis of the time and effort needed for recording and measuring the unit and its behavior."

So we know that a very wide range of chemical products have a very wide range of behavioral effects, but the correlation between potency of drug and potency of effect in the human being gives about the same size of validity coefficients that experienced investigators have found in psychological research on human personality—large enough to be encouraging, but too small to be smug about. And the so-called psychological effects—whatever it is that makes the placebo potent, or makes the enthusiasm of the physician potent—may well be as important as or more important than, the chemical structure of the compounds so many of us now take.

U. S. Navy Medical Neuropsychiatric Research Unit
WALTER L. WILKINS

Organizations. By JAMES G. MARCH and HERBERT A. SIMON with the collaboration of Harold Guetzkow. New York, John Wiley and Sons, Inc., 1958. Pp. xi, 262. $6.00.

Modern Organization Theory. Edited by MASON HAIRE. New York, John Wiley and Sons, Inc., 1959. Pp. x, 324. $7.75.

Though problems of organizational behavior have occupied philosophers for more than 25 centuries, it is only recently that these problems have been approached in a systematic, empirical way. The problems are very complex and require contributions from anthropology, economics, sociology as well as from psychology. The two books under review constitute important contributions to this field of knowledge.

March and Simon, two political scientists, and Guetzkow, a psychologist, bring together the scattered and diverse body of writing about organizations. The authors accomplish this as well as could be expected at the present stage of development of this field of knowledge. In a *tour de force* of erudition they define more than 200 variables and state several hundred propositions of the "Y is a function of X" type. Each of these propositions is justified with pertinent references to the literature. However, as might be expected in so young a field, some of these references are only suggestive of the proposed relationship. These intellectual pyrotechnics occur in only about 200 pages, yet the book reads as well as most review articles. The authors refer to more than 1000 titles from the literature.

The psychological assumptions that underlie this book are compatible with the theorizing of Tolman, Bruner and his associates and most particularly Newell, Shaw and Simon (*Psychol. Rev.*, 1958, *65*, 151-166). The book is divided into three parts. The first part deals with "classical organization theory"—F. W. Taylor (1911), Gilbreth and Gilbreth (1917), Fayol (1930), Gulick (1937), Mooney and Reiley (1939) and Urwick (1943). The authors discuss the inadequacies of this theory. The second part deals with models of bureaucracy developed by Merton (1940), Selznick (1949) and Gouldner (1954) and introduces a simple model of adaptive motivated behavior. This part concentrates on the relationships between productivity and a number of variables, including job satisfaction, and the amount of organizational control over these relationships is explored. An important chapter, in this part, deals with organizational conflict. The third part of the book is an extension of the model developed in the previous part. The authors examine critically the model of rational man used in economics and statistical decision theory and propose several modifications. They argue convincingly that planning and in-

novation in organizations follows laws that are parallel to the laws of human problem-solving. Hypotheses concerning the occurrence of innovation are derived from a knowledge of the human process of problem solving.

Thus, March, Simon and Guetzkow have made a major contribution to organizational theory, not in the sense that they have provided a final theory, but in the sense that they have defined the variables, sketched out the most likely relationships, and formulated their propositions so that they are amenable to empirical testing. However, much remains to be done: Adequate methods of measurement for most of the variables are still lacking; the framework that they provide has large gaps in terms of empirical evidence. Nevertheless, the book constitutes an important first step towards a theory of organizational behavior.

The second volume under review is edited by Mason Haire. It includes papers presented at a symposium held by the Foundation for Research on Human Behavior, in February of 1959. It brings together the contributions of a number of leading experts in economics, industrial administration, psychology and sociology. Haire, in the introduction to the volume, writes of the diversity of approach and of theoretical backgrounds represented in the volume and states that the papers are "the ragged leading edge of a wave of theoretical development," but admits that they "bear only the common imprint of a focus on the organization of productive human groups." He attempts to unify the papers by defining some of the issues of organization theory—the conflict between personality and organization, the structure of organizations, decision theory and organizational survival, etc.—and by showing how each of the papers fits into the issues, but in spite of this valiant effort the book does not have any unity. However, the reader who wishes to obtain a quick view of the landscape of organization theory and wishes to become familiar with some of the thinking of a number of the leading people in this area will find this book very useful.

E. Wight Bakke, in the following chapter of the book, discusses the concept of social organization and concludes that "a social organization is a continuing system of differentiated and coordinated human activities utilizing, transforming, and welding together a specific set of human, material, capital, ideational and natural resources into a unique problem-solving whole engaged in satisfying particular human needs in interaction with other systems of human activities and resources in its environment." This definition is elaborated, and the author's well known "Bonds of Organization" are defined and discussed. Bakkes framework is comprehensive and elegant. However, to date, very few empirical studies have made use of it, and like any other framework it can only be evaluated after its usefulness has been proven in the arena of data.

The third chapter, by R. M. Cyert and J. G. March, argues that organizations make decisions in the same sense in which individuals make decisions. The authors expand on the arguments presented by March and Simon (1958), and attempt to formulate a theory of organizational objectives. The theory utilizes many of the notions of the level of aspiration which have been developed from individual goal-setting behavior.

Anatol Rapoport, in the fourth chapter, describes a fascinating procedure for the laboratory study of group problem solving. The Electronically Operated Logical Test provides both great flexibility and the means for a quantitative measurement of important variables such as group performance. It is a significant contribution to the methodology of the study of problem-solving by groups.

The fifth chapter is by Chris Argyris and focuses on the conflict between individual and organization. Argyris attempts to list all the variables that are relevant in a given situation, but this procedure leads to propositions of the form "if X then A or B or C or . . . or Z." Such propositions are not as useful as propositions of the form "if X then Y." Argyris might argue that the latter kind is unobtainable at the present stage of development of organization theory; however, March and Simon (1958) have shown that it is possible to go somewhat further than the kinds of propositions used by Argyris. On the other hand, when it comes to a discussion of data, Argyris has much to offer. His attempt to develop self-actualization scores deserves close attention. His validation of these scores with turnover criteria seems very promising.

W. F. Whyte discusses interaction theory in the next chapter. He further develops the simple yet powerful theoretical scheme used by Chapple, Arensberg, Homans and others and applies it to organizations. He defines interactions, activities and sentiments and discusses their intercorrelations. He suggests how they are to be measured, and illustrates his points with a number of case studies.

The seventh chapter, by Rensis Likert, attempts a modification of the current theories of organization and management and stresses motivational variables. Likert uses economic, ego, security and curiosity and creativity motives. He documents his arguments with discussions of how supervisory behaviors relate to differences in the attitudes of workers concerning job relevant matters.

Robert Dubin, in the eighth chapter, discusses the stability of human organizations. The approach has much in common with the early work of Bavelas and Leavitt on communication nets, but it is expanded with notions from information theory.

The ninth chapter discusses the potential contributions of graph theory to organization theory. Dorwin Cartwright asserts that the mathematical theory of linear graphs can be of considerable help to the organization theorist in dealing with the structural properties of organizations; he presents some of the elements of this theory and discusses how it can be related to organization theory.

Chapter 10, by Mason Haire, presents some fascinating studies of organizational growth. It shows how data fit into a simple equation, $dN=(K-N/K)\cdot N\cdot \log_e R\cdot dt$, where N is the number of people in the organization, K is the limit on the size of the population imposed by the existence of limited environmental support, R is the rate of growth during the first three years of life of the organization, and t is time. R and K are associated with characteristics of the industry. Haire further suggests a framework for the study of the growth of staff and line.

In the final chapter, Jacob Marschak discusses efficient and viable organization forms. He explores these forms from the point of view of decision theory and shows how messages received from the outside world and from within the organization can be combined in a maximally effective way so that the actions of the organization will lead to the highest chance of survival. He recognizes that his "model" differs from "reality" but argues eloquently for the need for this kind of theorizing.

Here then is a book that brings together ten points of view on how to study organizations. Each of the chapters can be read independently and the points of view are only slightly related to each other. Most readers will find promise in many of the approaches that are proposed. Which of the approaches, or combination of approaches, will lead to the greatest advance in organization theory remains to be seen.

Graduate programs in social or industrial psychology may benefit from the inclusion of a course in organization theory. The materials for such a course are now available: Text: March and Simon (1958). Readings: Haire (1959).

University of Illinois HARRY C. TRIANDIS

Patterns of Child Rearing. By R. R. SEARS, E. E. MACCOBY, and HARRY LEVIN. Evanston, Illinois, Row Peterson and Co., 1957. Pp. VII, 549. $5.25.

This volume is a report on an investigation that was made to answer three questions: (1) What is the current prevalence of selected child rearing practices? (2) What are the differential effects on the children of different approaches to child rearing? (3) What leads a mother to use one child-training method rather than another?

The report presents considerable data on these questions from a selected sample of 379 mothers. The sample represents the residual of 640 cases remaining after a variety of exclusions, for example: 38 cases were dropped because the parents were foreign born, 79 cases were excluded to produce the desired sample composition with respect to social class and ordinal position. There were 38 refusals (8% of the 481 contacted), a group distinguished from the 24 who "did not have time to be interviewed." Regarding the refusals, the authors report that "they worry us . . . and we suspect that they may have represented extremes in one way or another" (p. 25). Regarding the 24 who lacked time to be interviewed, the authors report parenthetically, that "we listened to their reasons and believed them!" (p. 25). Perhaps it would have been more appropriate to believe them but still classify them as refusals.

These were the approaches to and attitudes about sampling problems reported in a book the first purpose of which is to report on the *prevalence* of child rearing practices. A statement of prevalence requires the identification of a population and a period in time. The period of time is well specified, but the sample cannot be considered to be representative of any population at all—not even of the cities studied, because of the exclusions. In the words of the authors, "This book will report a study of child rearing in the United States as one group of 379 was practicing that art at the close of the first half of the twentieth century" (p. 27). If the words "in the United States" had been omitted or specifically modified, the statement would be more nearly precise.

With respect to their second question—the effects on the child of differences in child rearing practices—the authors are careful to point out that both the "cause" and "effect" for any one child were measured by one interview with the same mother, and that the measures lack methodological independence. A major finding is that severe punishment was associated with both dependent and aggressive children. This is an important finding. The authors point out that their data do not show whether the child's aggression induced the punishment or the punishment induced the aggression. They nevertheless summarize: "Our evaluation of punishment is that it is ineffectual over the long term as a technique for eliminating the kind of behavior toward which it is directed" (p. 484). In so far as the reader can re-state the findings as showing simply the existence of a relationship, the material is highly significant and of considerable interest as correlational research.

The authors give a detailed account of their method of treating their data. Each of the 379 of the interviews recorded verbatim was rated twice on 188 dimensions. The raters were 10 advanced graduate students, employed in rotated pairs. Of the total 188 dimensions 23 appeared to relate to child behavior. Some allowed for the

possibility of objective reporting (*e.g.* how often child wets his bed), and some almost precluded objective reporting (*e.g.* how much of a problem is obedience).

Of the 165 dimensions of parental behavior or characteristics, 44 were subjected to factor analysis. The yield was seven factors. The 23 dimensions of child behavior were not subjected to factor analysis. This differential treatment is not explained.

The authors conclude that "the most prominent trait that influenced mothers' child rearing practices" (p. 472) was that represented by Factor A, *i.e. Permissiveness-strictness.* Despite the fact that it is the "most prominent" trait, they make no attempt to score it; nor, indeed, do they report its prevalence or its effect on child behavior. Perhaps they were concerned with the need for replication of such a factor analysis. They select, nevertheless, a "pattern" of *love-oriented vs. object-oriented* techniques of discipline, and list three dimensions for each, stating that the pattern "does not come from the factor analysis, but from a study of the intercorrelations among methods of discipline" (p. 477). In this case, as in others, the authors did not feel a need for the replication of correlational findings prior to interpretation.

Their work led the authors into controversy with Davis and Havighurst about differences in social-classes permissiveness in child rearing. They found that mothers of the middle-class are more permissive than mothers of the lower-class.

The factor seen as having the most significant influence on child behavior was *Warmth vs. Coldness.* This time Factor C identifies the "trait." "Warmth" was defined by ratings reflecting (1) much affectionate interaction with the baby; (2) high affectionate demonstrativeness toward the child; (3) praise for good table manners; and (6) use of reasoning as a method of training. The mothers who reported behavior which was perceived by the raters as fitting the six foregoing characteristics were also, generally, the mothers who reported few feeding problems, no persistent bed-wetting, low aggression, no upset during toilet training, and ready "conscience development." Again, as the authors mention, only the existence of the association is established, the maternal behavior may have induced the child behavior, or *vice versa*—or there may have been common induction of both, or there may have been a circular or "chain" reaction—in any case the association will no doubt have considerable influence on later work, if the assessments can be repeated.

The authors report an additional interpretation of great potential importance: that intra-family aggression takes on a heirarchy—the individual acts aggressive mostly toward the next highest available person in the family power structure. Only and oldest children were more aggressive toward their parents than would have been predicted from child rearing practices. Middle and youngest children were more aggressive toward their older siblings.

In general, despite its limits, this is likely to be a most influential book in the field of child development.

St. Louis County Department of Health. JOHN C. GLIDEWELL

The Life and Work of Sigmund Freud, Vol. III. By ERNEST JONES. New York, Basic Books, 1957. Pp. xvi, 537. $7.50.

A comparison of the three successive volumes of this widely-hailed, authorized Life of Sigmund Freud, now completed, shows a progressive decrease of biographical material and a concomitant increase of doctrinal exposition. As biography the first volume is clearly superior to the other two for this reason alone, but one has the

further impression that the biographer gradually retreated from his assumed role to the more comfortable one of discipleship.

Reflecting on this fact, among others, one reaches the conclusion that this biography stands alone as a paradoxical expression of allegiance: it is written as it would have been had its subject never lived; it applies none of the discoveries of Freud to the illumination of his life and work. Conceive of Francis Darwin compiling the official life of his father Charles and writing of him as having had an immortal soul added to his animal body by a divine power at some prenatal moment; conceive of Rheticus, writing the only contemporary biography, now irretrievably lost, of Copernicus and describing the day of his subject's birth or death in the terms of a Ptolemaic astronomy; conceive of the Synoptic disciples composing the biography of their Master in pagan images, invoking upon Him the blessings of Astarte or Osiris—and you will have put yourself in the proper frame of mind for reading Jones on Freud. Moreover, you are left in no doubt by this biographer as to his intention: he asserts that he will not psychonanalyze Freud, implying that to do so would be presumptuous; so, instead, with a disarming show of frankness, a humorous quip, or a time-honored aphorism, he repeatedly rings down between his subject and his audience a curtain of rationalization.

Yet Jones can be flagrantly inconsistent in this regard. A striking example is in his discussion of the death instinct—a concept that he, in common with many other Freudians, rejects. To justify the rejection he delves at length into Freud's unconscious motivation for this mistaken concept, albeit such use of psychoanalysis comes dangerously close to the *ad personam* technique that the logic books condemn.

One of the pervasively disturbing elements of these volumes is the almost consistently derogatory picture of Freud's successive associates and rivals—which they rapidly became in that association. Jones willingly sacrifices them one by one in justification of Freud, disposing thus of Breuer as reactionary and prudish, of Fliess as "pathological" in personality, of Ferenczi as suffering an unfortunate decline in mental health. Even Havelock Ellis, not an associate but a contemporary from whom Freud heavily borrowed such concepts as autoerotism, is dismissed as "jealous" of Freud's greater success. That Jones in these instances should take the point of view of his biographee is not surprising, but one somehow feels that, especially at these foci, the identification is closer than a balanced judgment would permit.

For the psychologist one of the most pertinent sections of the third volume is Chapter 9, Lay Analysis. Recognizing that this professional issue is still unsolved, Jones takes what he calls a "midway position." Freud was, on the other hand, univocally in favor of lay analysis. This chapter contains valuable historical data as well as cogent arguments *pro* and *con* that professional psychologists, concerned with legislation for licensing, etc., could profitably use.

Despite its shortcomings, which are destined to become more obvious with the passage of time and hence more readily discounted, this book is a storehouse of factual information for which any future biographer will be grateful—and there will, of course, be such biographers. The life of Freud was an "experiment of nature"—an experiment of human nature—for which few if any precedents exist in recorded form. One may question now or in the future the validity of the "unconscious" but Freud's superconsciousness of himself and his world over a period of eighty-three years, all of them more than less available, is unequalled for the understanding of normal and abnormal personality in its furthest reaches.

It is in keeping with the uncanny quality of his other pronouncements that Freud should have delivered in advance the most trenchant characterization of this biography. Jones narrates that Freud was in his later life alarmed to learn that his friend Arnold Zweig was planning to write a biography of him. "He firmly forbade him to, telling him he had far more useful things to write. Freud's views on biographical writing were certainly extreme, since he added: 'Whoever undertakes to write biography binds himself to lying, to concealment, to hypocrisy, to flummery and even to hiding his own lack of understanding, since biographical material is not to be had and if it were it could not be used. Truth is not accessible; mankind does not deserve it, and wasn't Prince Hamlet right when he asked who would escape a whipping if he had his deserts?' " Unabashed Jones goes on: "And yet I continue with my task in the fact of these terrible dicta; I feel sure that Freud would have been surprised to find that one could get nearer to the truth about himself than he imagined possible." The question is an open one.

Washington University, St. Louis SAUL ROSENZWEIG

Educational Psychology and Children. By R. LOVELL. New York, Philosophical Library, 1959. Pp. 272. $6.00.

In his foreword to the book, P. E. Vernon states that "we (in England) have found it difficult to progress much beyond the position to which they (William McDougal and Charles Spearman) brought us in the 1920s." In attempting to bring an up-to-date psychology to teachers, the author includes some 17 chapters on topics generally discussed in American books of Educational Psychology. He also devotes three chapters to the characteristics of the pre-school child, the primary school child, and the adolescent. The book is documented with a considerable number of research studies. At the end of most chapters the author states the implications of the quoted research for children's development and learning.

Unfortunately the topics selected are not adequately covered. This is partly the result of including the necessary basic psychology in the book. It is impossible to present general and educational psychology adequately in such a short text. This superficial coverage is not, however, entirely the result of the book's limited size. Some topics have been given far greater emphasis than their importance would warrant. For example, three pages are devoted to a discussion of coaching for intelligence tests. Coaching is of concern in Britain because of the competitive examination used to select children for grammar school education. It seems, however, unnecessary to devote more than a few lines to a practice so obviously wrong.

Although it would be impossible to include all topics relating to educational psychology, this book has some glaring omissions. The author mentions Thurstone's group abilities, but he does not discuss differential aptitude testing. He discusses stimulus-response-associationism and field-cognition learning theories, yet he makes no mention of Skinnerian conditioning nor of Lewin's topological theory. He explains growth curves but he includes nothing about the Fels Institute and Harvard growth studies or the Berkeley studies on growth and decline.

In addition to omissions, he has, in at least one instance, used research that has been generally discredited, to support an argument. I refer to his use of the Bernadine Schmidt article that appeared in the *Psychological Monographs* in 1946 to show the impact of improved emotional adjustment on measured intelligence.

There appears to be no plan for the text. The book as a whole is not well integrated. Chapters are discrete units and no attempt is made to relate them to one another. We seriously question whether this text will contribute to the knowledge of the beginning teacher or to the progress of educational psychology in the United States.

Cornell University M. D. GLOCK

Ten Thousand Careers. By ROBERT L. THORNDIKE and ELIZABETH HAGAN. New York, John Wiley & Sons, Inc., 1959. Pp. vii, 346. $8.50.

This book is a detailed report of a study, financed by the Grant Foundation, of a group of men who were Aviation Cadets during World War II. The subjects, who had taken a battery of aptitude tests and had filled out a bibliographical data blank in 1943, were surveyed by questionary in 1955 and 1956 to get information about the vocational status and success of each man. Thorndike and Hagan present a careful analysis of the test-scores and biographical items in relation to the careers of the 10,000 men from whom it was possible to obtain follow-up data. General findings include the following: The occupational groups studied differ significantly from one another in respect to certain mean test-scores and variables in personal background. Aptitude profiles for some groups are noticeably uneven; for others, they are relatively flat. Within any given occupational group, there is a wide range of scores for each aptitude. There is no convincing evidence that relative success (as defined) within any occupation could have been predicted from either the test-scores or the biographical information.

The authors take pains to state, more than once, that their project had several limitations of consequence. These are described candidly by Thorndike and Hagan, who refer to their investigation as "less than perfect as a definitive study of the effectiveness of aptitudes as predictors of long-range occupational choice or success" (p. 48). The publishers are more pretentious; they make reference to the book in these words: "It promises to be the (*sic*!) major source of data for anyone teaching or working in the field of vocational guidance or personnel selection" (jacket). This statement is an excellent illustration either of wishful thinking or of hyperbolic writing. The book should prove to be useful, yes, but its import and value are circumscribed by the limitations of the research on which it is based.

This reviewer wishes that the authors had taken the trouble to explain the differences between long-range occupational prediction and the processes of vocational counseling and employee selection. Their failure to do so, together with the proclivity of members of the fourth estate to accentuate the negative, has resulted in some garbled popular reviews, and in newspaper headlines such as these: Aptitude Testing Not So Apt; Aptitude Testing Is Found False Profit On Jobs.

The writing is clear and interesting. The book has an attractive format, and it contains a large number of informative figures and tables.

Cornell University A. GORDON NELSON

Private Practice in Clinical Psychology. By THEODORE H. BLAU. New York, Appleton-Century-Crofts, Inc., 1959. Pp. x, 184. $3.00.

In recent years there has been a substantial increase in the practice of clinical psychology. A book to help people entering or preparing to enter the field fulfills a professional need. To satisfy this, Dr. Blau has, in the light of six years experi-

ence, written this book about clinical practice. For this purpose he has selected materials which he presents in four parts, captioned: Preparation for Practice, Clinical Functions and Procedures, Professional Liaison and Responsibility, and Additional Considerations. Illustrative topics considered include: The Candidate for Clinical Practice, The Clinical Setting, Organizing the Practice, Patients and Their Problems, Evaluation, Therapeutic Functions, Professional Relationships, Ethical Standards, Fees, The Clinical Psychologist and the Legal Profession, Research in Private Practice, and Group Private Practice. Also included is an Appendix with directory material and sample forms, form letters, and procedures.

The author begins the book with the statement that, "Before a detailed analysis of the functions of clinical psychologists in practice can be made, the background of the clinician must be considered." This to him means a consideration of the education, experience, a definition of private practice, a description of the American Board of Examiners in Professional Psychology and the American Board for Psychological Services. It does not include a historical perspective of the field. It is the author's impression that, "Nowhere in the professional literature has there appeared, to date, a critical yet constructive analysis of private practice." A more correct statement, however, is that the author limits his coverage of references, almost exclusively, to articles appearing in *The Ameriacn Psychologist*. There is no evidence, in this book, of familiarity with many references that are available in considering private practice. the nature of the problems, and the place of private practice in professional psychology.

There probably is a need for a book on private practice, but the content of this book fails to satisfy this need for most professionals. The most that can be said of it is that perhaps it can be of some help for the inexperienced who are not resourceful enough to prepare their own forms and write their own letters.

Human Relations Research Foundation — HYMAN MELTZER
St. Louis 8, Missouri

Instinct in Man in the Light of Recent Work in Comparative Psychology. By RONALD FLETCHER. New York, International Universities Press, 1957. Pp. 344. $7.50.

The author has attempted a critical survey of theories of instinct with special reference to man, from Darwin to the present day. In Part I he covers the "earlier views," including, among others, Darwin, William James, Lloyd Morgan, Hobhouse, and McDougall, and then offers a criticism. Part II is devoted to the work in the newly titled field Ethology—the comparative study of animal behavior in biology—and to the treatment of instincts in psychoanalysis. Part III is a synthesis based on a critical analysis of the early views, of the contributions of ethology, and of the doctrine of Freud. This section concludes with a seven-page table of what are called Instincts Proper (primary impulses), General Instintictive Tendencies (ego-tendencies), and Secondary Impulses; and a final chapter on some implications of the whole inquiry. Since the author is a Lecturer in Sociology (at the University of London), it is not surprising that the upshot lies for him in the meaning of the instinct concept for social psychology and sociological theory. One is reminded of McDougall's initial presentation of his theory in his 1908 *Social Psychology*.

This work is a useful compendium, within its scope, and contains some shrewd

observations. Despite its critical aim it suffers, however, from being restricted to the contributions of biologists (ethology) and psychoanalysts; there is in it practically no recognition of the related and, in some respects, more rigorous contributions of comparative psychologists in the stricter sense of that term. The Bibliography lists essentially nothing by the leading animal *psychologists* and for that reason the subtitle of the book is misleading. More significantly, this omission limits the monograph to an attempt at correlating the findings of workers like Tinbergen and Lorenz with the pronouncements of Freud. The more operationally defined concepts of learning theory and the (sometimes) better controlled observations of the animal psychologist fall outside the author's ken and hence cannot enter into his critical formulations. For the psychologist the volume is therefore more suggestive than definitive; it will be more acceptable to some personality theorists than to most experimental psychologists.

Washington University, St. Louis
SAUL ROSENZWEIG

Sigmund Freud and the Jewish Mystical Tradition. By DAVID BAKAN. Princeton, New Jersey, Van Nostrand, 1958. Pp. XIX, 326. $5.50.

The major aim of this book is to show that the intellectual origins of psychoanalysis can be found in Jewish mysticism. Freudian ideas are considered as a latter development of cabalistic thought ripened and mellowed by contact with the scientific attitude.

Two lines of supporting evidence are developed. One deals with the ethnic-historical aspects of the thesis, the other with more direct parallels between the content of psychoanalysis and cabala. Central to the first is an event of the 17th Century which the author believes to have profoundly influenced the future course of Jewish thought; namely, the appearance of a false Messiah, Sabbatai Zevi, who promised an oppressed people deliverance. His significance for later Jewish thought was to implant within the culture the messianic ideal, the will to fight against orthodoxy, and the wish for apostasy. These concepts are related in the narrative to the development of Freud's life and his ideas. An analysis of Freud's identification with the figure of Moses plays an important part in this story.

Dream theory and the fundamental importance of sexuality are the content issues around which the similarity between the two systems of thought are drawn. Many of Freud's more important discoveries are traced to ideas present in the *Zohar,* the most influential book of Jewish mysticism.

The question of Freud's awareness of his intellectual debt, a matter of considerable importance for the thesis, is handled by the author in the following manner. He states that Freud like others in the mystical tradition could have dissembled and points out this tendency in his published work. The alternative is also offered that the ideas in question could have reached Freud through the milieu of his own sub-culture.

The book is distinguished for the original position it presents and the ingenuity with which the evidence is marshalled. The evidence itself however, mainly interpretive in nature, is not likely to be readily accepted by students of Freudian history. It is a highly provocative speculation which invites equally speculative rebuttal.

IRVING E. ALEXANDER
National Institute of Mental Health
Bethesda, Maryland

Psychopathology: A Source Book. By CHARLES F. REED, IRVING E. ALEXANDER and SILVAN S. TOMKINS (Editors). Cambridge, Mass., Harvard Univ. Press, 1958. Pp. xii, 803. $12.50.

This volume is a collection of 46 articles selected as representing mainly recent trends in research on psychopathology. It was not the intention of the editors to produce a reference source representative of the field of psychopathology in general.

Nearly all the articles selected were originally published during the period 1952-57 and almost all are important papers of high quality. The majority are research reports, though a few are theoretical, such as Redl's paper on ego disturbances, Selye's *Stress and Disease*, and Calvin Hall's theory of dream symbols. The book has an effective index, both of authors and subject matter.

Five broad areas of investigation are represented. Part I deals with "Psychopathology and Early Experience" and contains articles ranging from early infantile autism through early experience studies in the Hebb tradition. For this reviewer this is the strongest section of the book. Part II, entitled "Psychosomatic Disorders and Neurosis," is long on psychosomatics but short on neurosis. As a consequence it and Part III, "Schizophrenic Psychoses," struck the reviewer as the weakest sections. Better than half the articles included in the latter deal primarily with treatment processes, rather a high proportion for a text in psychopathology. On the other hand, Part IV includes a most satisfying total of eight papers concerned with "Somatic Factors in Psychopathology." Most of these have reference either to phenomena of brain injury or to biochemical aspects of schizophrenia and "experimental psychosis." Finally, Part V, "Psychopathology and the Social Context," is an adequate representation of recent studies in epidemiology and other socio-cultural factors related to psychopathology.

This is a worthwhile book. With the qualifications noted above, what it does, it does well. It should prove particularly useful as a library reference, especially where periodical resources are sparse.

Washington University RAYMOND G. HUNT

Language and Psychology. By SAMUEL REISS. New York, Philosophical Library, 1959. Pp. 299.

With the currently mounting interest in the psychology of language there should be an eager welcome for a study of the process of vocabularly creation "as a means or tool for obtaining new psychological insights into the mode of the human thought process." The psychologist will, however, be disappointed; and the linguist will be irritated. The 'theory' presented bears no relation to recent developments in either psychology or linguistics; except for a single footnote reference to Freud, and another to Cherry, the only sources mentioned are the author's own previous books; and the only direct grappling with alternative theories is to be found in a summary rejection of positivism and information-theory. There is no table of contents, no bibliography, and only a very meager index.

The theory presented reduces linguistics to psychology. It emphasizes the 'natural' origin of words from action sounds nad the almost limitless elaboration of connotative meaning through metaphorical association. Most linguists will not like this. Most psychologists will wince at the kind of psychology to which linguistics is being reduced, a thoroughly old-fashioned 'mentalistic' psychology. At least for the psychol-

ogists, however, it should have some value, for it provides 45 elaborately prepared lists of English words and their definitions, plus a similar list of 15 Japanese words, each list illustrating some aspect of the general theory. Whether or not the theory is actually substantiated by the lists will have to be decided by the linguists, but psychologists who have become interested in the physiognomy of language may find them of more than passing interest.

Cornell University ROBERT B. MACLEOD

Personality and Persuasibility. Edited by IRVING L. JANIS and CARL I. HOVLAND, New Haven, Conn., Yale University Press, 1959. Pp. xi, 333. $5.00.

This volume, the second in a projected series of studies on attitude and communication, is a superior specimen of a genre increasingly common: a collection of papers by a number of different authors dealing with various aspects of the same subject. It is rather better than most productions of this kind because it was planned all as a piece, and carried to completion as a single project.

Briefly, the book is concerned with an investigation of "factors which underlie responsiveness to one or another form of social influence." The central theme is carefully delineated by the editors in a first chapter 'overview.' This is followed by a series of research papers dealing with some of the factors involved in persuasibility. Finally, the editors have added a well-written summary and theoretical discussion. On the basis of their research, the authors conclude that people differ markedly in "general persuasibility," a behavioral dimension essentially independent of specific environmental condition. Persuasibility appears to be a function of, among other things, age, sex, and various personality characteristics. At present, it is not demonstrably related to intelligence. One interesting set of findings dealt with the effects of cultural role on persuasibility. It seems that, in adults, personality factors are significantly related to persuasibility only in men; women being persuasible regardless of individual turns of personality. On the other hand, young girls (age 6 yr. or so) do not differ from boys of the same age in their responses to social influence.

Obviously, such a book as this has value for persons closely interested in communication. Those doing research on persuasibility will doubtless find the detailed reports of supporting experiments useful in formulating their own plans. The more general reader, however, can find out all he needs to know from reading the first and last two chapters, which contain all the basic material in the book.

N. G. BURTON
San Antonio, Texas

Social and Moral Attitudes. By S. BRAHMACHARI. Calcutta, India, Orient-Longmans Private Ltd., 1954. Pp. 113. 5 rupees.

This work is concerned with an examination of the interrelationships between attitudes, personality type, and conditions of upbringing of 120 educated adult subjects. Data were obtained from 4 customized rating scales plus the Woodworth Neurotic Inventory. Average responses and various intercorrelations were computed. Some effort was also made to ascertain the sources of the attitudes and values studied. All results are interpreted according to psychoanalytic theory, with the author tending to form his arguments as homely discourses between ego, superego, and id.

Apparently this study was conducted about thirty years ago, the latest work cited in the bibliography having appeared in 1929. According to a foreword by J. C. Flugel, it was for many years an unpublished doctoral thesis on file in the University of London Library. It was finally published (still in standard thesis form) only after frequent borrowings by students all over the world wore out all the library copies. Under such circumstances as these, it is hardly fitting to criticize the research described as though it were of recent origin. Suffice to say that, in the light of thirty years' development, this study's principal value is historical. Because it is a 'pioneer work,' however, it will doubtless continue to serve as a valuable reference for those planning research in the area of attitude and value study.

San Antonio, Texas N. G. BURTON

Gestalt Psychology. By WOLFGANG KÖHLER. Rev. Ed., New York, New American Library of World Literature, Inc., 1959. Pp. 224. $0.50.

This edition of Köhler's *Gestalt Psychology* is a reprint of the hard-cover edition published by Horace Liveright, Inc., in 1929, which was reviewed by Professor Harry Helson in this JOURNAL in 1930 (Vol. 42, pp. 657-662). As the first edition has long been out of print, this inexpensive, paper-covered, Mentor Book is very welcome. A student will no longer need to seek the loan of his professor's copy because the library copy is in use and unavailable. It is a pity that more of the important books in psychology are not reprinted in Mentor editions.

Though the new edition was extensively rewritten—the titles of 4 of the 10 chapters were slightly altered; a number of paragraphs in the various chapters were deleted and others added; and many sentences in the paragraphs retained were, for the sake of style and clarity, rewritten—the outline and substance of the book, its vigor and point of view, are unchanged. Helson's excellent review is, consequently, as adequate for this edition as it was for the first.

In addition to changes in format, the bibliographies, which appeared at the end of every chapter in the first edition, have been revised and transferred, under the various chapter headings, to an appendix at the end of the book. A thoroughly revised index, subject matter and authors combined, closes the volume.

K. M. D.

The Dissociation of a Personality: A Biographical Study in Abnormal Psychology. By MORTON PRINCE. Second edition, reprint. New York, Longmans, Green and Co., 1957. Pp. x, 575. $5.00.

This reissue of the second edition (1908) of Prince's masterpiece, now seldom read, about the multiple personality "Miss Beauchamp" is a timely reminder of his classic contribution to psychopathology. Prince rightly feared that his work might be eclipsed by the stronger appeal of Freudian psychoanalysis. It was, in fact, part of his motivation in publishing this extended "biographical study in abnormal psychology" to get some of the public's attention—a "ruse" to which he himself confessed in the Preface to his last book (1929). The rich detail of the volume has today, however, an intrinsic value quite apart from the motivation that produced it. With the current revival of interest in hypnosis even in psychoanalytic circles (witness Gill and Brenman's *Hypnosis and Related States,* 1959) and the recognition, by at least some, that Freud's "rejection" of hypnosis may itself have

had unconscious motives, this new printing of the *Dissociation* makes the book again readily available, as it should be. The first edition of this book was reviewed in this JOURNAL (Vol. 17, pp. 280-284) by Professor Joseph Jastrow in 1906.

Washington University, St. Louis, Missouri — SAUL ROSENZWEIG

Course in General Linguistics. By FERDINAND DE SAUSSURE. Edited by Charles Bally and Albert Sechehaye. Translated from the third French edition by Wade Baskin. New York, Philosophical Library, 1959. Pp. xvi, 240. $6.00.

Since 1915, the *Cours de Linguistique Generale*, assembled from lecture notes by two of de Saussure's pupils, has been a standard reference work for students of linguistic theory. De Saussure is best known to psychologists for his distinction between diachronic and synchronic linguistics, a distinction which is vital for any psychology of language. Psychologists as well as linguists will consequently welcome this competent and well edited translation.

Cornell University — ROBERT B. MACLEOD

Money and Emotional Conflicts. By EDMUND BERGLER. New York, Pageant Books, Inc., 1959. Pp. xiii, 269. $3.95.

This is the second edition of a book which first appeared in 1951, reissued, according to the foreword, by continued demand. It does not pretend to make unique contribution to personality theory. Rather, its primary purpose is to present an established point of view (psychoanalysis) dealing with a particular type of adjustment problem (conflicts centering around money) to a reasonably intelligent lay audience. Within these self-set limits it is adequate and makes interesting reading. It is not technical or thorough enough to have value as a reference work for clinicians or psychiatrists.

San Antonio, Texas — N. G. BURTON

Psychiatry in the British Army in the Second World War. By ROBERT H. AHRENFELDT. New York and London, Columbia University Press and Routledge and Kegan Paul, 1958. Pp. xiii, 312. $6.00.

This is a revision and abridgment of the semi-official history of British Army psychiatry compiled in 1948. The principal areas of military effort of the psychiatrists are discussed in chapters dealing with selection, mental defect and slow learning, disciplinary problems, training and morale, rehabilitation of prisoners of war, psychiatry in forward units, and disposal of psychiatric cases.

As with the histories of other military units—the WAC, for example—one can see many instances in which technical insights and skills were not used in the way they might have been. British psychiatry had a strong skeptic in Winston Churchill. This record of trials and triumphs has lessons for management of personnel, civil as as well as military, but it is just these lessons that peace-time administrators and practitioners are inclined to ignore.

St. Louis University — WALTER L. WILKINS

Handbook of Philosophy. By M. H. BRIGGS. New York, Philosophical Library, 1959. Pp. 214. $4.75.

Briggs' *Handbook* gives a quite elementary compilation and explanation of philosophical terms through simple yet competent definitions. It is of interest to the

psychologist because it includes many terms which exist in the conjuncts of the two disciplines as well as within psychology itself.

New York WARREN R. STONE

Dictionary of Personnel and Industrial Relations. By ESTHER R. BECKER. New York, Philosophical Library, 1958. Pp. 366. $10.00.

Becker's *Dictionary* contains 2468 entries covering industrial organization, labor and union (management and personnel) activities, industrial welfare programs, and wage determination, plus a section on research in personnel and industrial relations at universities. It is a complete coverage of current usage which should be of considerable value to industrial psychologists.

New York WARREN R. STONE

The O-Structure: An Introduction to Psychophysical Cosmology. By C. C. L. GREGORY and ANITA KOHSEN. Church Crookham, Hampshire, Institute for the the Study of Mental Images, 1959, Pp. 147, 21s.

'O-structure' is the name given to an informational model intended to describe every aspect of the world as part of a single system. The primary purpose of devising such a model is to avoid a dualism of physical and mental events, and the discussion specifically covers phenomena ranging from material objects to personality abnormalities.

Such a volume is to be commended (if not recommended) for its concerted effort to treat all the phases of the universe under a common rubric. By virtue of the fact that it does deal with the natural order in general, however, it becomes a philosophical work rather than a psychological one, and beyond demonstrating the mathematical and scientific feasibility of an absolute monism (a concept not previously unknown among psychologists) it contributes nothing particularly to the field of psychology.

San Antonio, Texas N. G. BURTON

Now or Never: the Promise of the Middle Years. By Smiley Blanton with Arthur Gordon. Englewood Cliffs, N.J., Prentice-Hall, Inc., 1959, Pp. 273, $4.95.

This is a sort of do-it-yourself mental hygiene text for persons between the ages of 35 and 60. Written in a brisk and informal style, it dispenses practical advice on problems characteristic of people who are mature but not yet elderly, *e.g.*, coping with teen-agers, attaining professional advancement, dealing with aged parents, keeping weight down, etc.

This book is intended primarily for laymen and, not surprisingly, offers little in the way of original psychological or medical thought. Its theoretical orientation is a typical blend of the views of Sigmund Freud and Rabbi Ben Ezra which the psychologically sophisticated reader could predict without ever opening the cover. It does, however, represent a real contribution to psychology simply because it deals with that phase of human development which has received the least notice from psychologists. The copious writings on childhood and adolescence and the current rage for studies of senility tend to overshadow the fact that the longest, most productive, *and* most problem-fraught phase of existence lies between the extremes of youth and old age. *Now or Never,* sketchy and pedestrian though it is, does at least call attention to the importance of the "middle years" and the need for further study of the phenomena they encompass.

San Antonio, Texas N. G. BURTON

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXXIII JUNE, 1960 No. 2

PERCEPTUAL ORGANIZATION IN SERIAL ROTE-LEARNING

By SOLOMON E. ASCH, Swarthmore College, JOHN HAY, Cornell University, and RHEA MENDOZA DIAMOND, Harvard University

In the study of serial rote-learning, temporal conditions always have held the center of interest. The method of investigation, which has varied little in essentials from the time of Ebbinghaus, illustrates this concern. The usual task requires the mastery of a set of discrete items, which appear repeatedly in a fixed sequence with strict control of temporal relations. The sole principle of order of such a series is temporal.

When one studies phenomena that succeed each other, the concentration upon temporal relations is appropriate and inevitable. At the same time, the usual procedure includes more particular features and assumptions that are deserving of scrutiny. A sheer temporal series is notable both for what it contains and excludes. In the first place, it excludes all other types of series (one of which we shall consider in this paper). In the second place, the temporal orders employed in serial learning have been of a special kind. They possess a minimum of structure, being uniform and monotonous.

The choice of procedure was dictated not only by considerations of simplicity, but by the assumption that it revealed the essential features of serial learning. Investigators have as a rule attempted to deduce the various phenomena of serial learning from associations between pairs of terms. They have treated the data of serial learning as the exclusive resultants of multiple dyadic associations, which in turn were referred to temporal conditions. Given this starting point, a uniform temporal series of discrete items does, indeed, recommend itself as most appropriate. This direction of thinking is notable for the assumption that a series *as such* introduces no proper-

* Received for publication May 19, 1959. This work was supported by a grant-in-aid from the National Science Foundation. We are grateful to the Department of Social Relations of Harvard University for its generosity in making laboratory facilities available for some of the experiments.

ties of its own that may be relevant to the course of learning. The properties of a series are, according to current interpretations, the aggregates of the associations included in the series. The use of bare temporal orders makes it difficult to detect conflicting data, thus apparently supporting the initial assumptions.

The following examples may serve as an illustration. A sequence may be described in terms of properties such as 'beginning,' 'end,' 'before and after.' The barest sequence, such as that of a set of nonsense-syllables, the members of which are discrete and arbitrarily ordered, possesses these (and perhaps other) minimal properties. Of every series it can be said that a given member is the first, and another the last; given a set of successive members, *a, b,* and *c,* we can say that "*b* is between *a* and *c.*" The mastery of a sequence can be described in terms of the progressively more accurate establishment of such relations. At this point, the following possibilities open up for thought. Serial learning may demand a direct representation of, and reference to, such properties of a series; or, the mastery of a series may be mediated by operations that do not require the noting and use of these relations. The issue is a factual one. Our ability to formulate the relations of a series analytically does not, of course, decide what the mediating operations are; for evidence we must go to the data themeselves.

Associative and, particularly, more recent stimulus-response accounts have consistently adopted the latter alternative. They hold that "*b* is between *a* and *c*" can best be described in learning terms as "*b* is the response to *a* and the stimulus to *c.*" The intent of such a formulation is to account for accurate ordering in terms of concrete learning operations that leave no place for a direct reference by the learner to the properties of a series.

Interpretations of the curve of serial position are a further example of the same point. Accounts such as those of Woodworth[1] and of Foucault,[2] or those of Lepley[3] and of Hull,[4] are at one in the following respect: They propose to derive the curve exclusively from the cumulative effects of individual associations. These formulations fail to consider the possibility that *S*'s discrimination of the boundaries of a series (and perhaps of its extent) also may contribute to the formation of the serial gradient.

We propose to examine the alternative possibility, namely, that serial learning may require the locating of an individual item with respect to the series as such. The direct noting and use of such relations as 'between,' 'beginning,' and 'end,' may be a condition of serial learning; and progress in learning may consist in part in the refinement of these distinctions.

Object: The investigations to be reported here grow out of the preceding question. Our object was to trace the effects of the properties of serial orders on the course of rote-learning. To this end we constructed a task that introduced clearly defined serial properties. The task called for the mastery, by the method of anticipation, of a series of nonsense-syllables.

---

[1] R. S. Woodworth, *Experimental Psychology*, 1938, 22-23.

[2] Marcel Foucault, Les inhibitions internes de fixation, *Année psychol.*, 29, 1928, 92-112.

[3] W. M. Lepley, Serial reactions considered as conditioned reactions, *Psychol. Monogr.*, 46, 1934, No. 205, 40-45.

[4] C. L. Hull, *et al.*, *Mathematico-Deductive Theory of Rote Learning* ,1940, 110, 175-177.

Our sole departure from the traditional technique was to impart to the series a spatial distribution. The syllables, which appeared singly and in succession, were exposed along a visual form or grouping that was constantly before the learner. Each member of the series had a particular and constant position within this spatial order on successive trials. This procedure established a *spatio-temporal* order, which permitted the introduction into the series of a range of perceptual conditions excluded by an exclusively temporal order.

The problem and the proposed procedure make contact with one of the first concerns in the experimental study of memory. It came early to the attention of investigators that *S*s learning a disconnected, uniform series attempt to impose an organization upon it by grouping the members, often by imparting a rhythm. Müller and Schumann were the first to demonstrate that rhythmic grouping of a list has a decided effect on learning.[5] Employing the method of derived lists, they demonstrated that syllables from the same rhythmic unit were associated more readily than syllables from adjacent units. Witasek similarly found that syllables which preserved their original place in a new rhythmic grouping were more easily associated than syllables whose rhythmic place was altered.[6] These experiments established that grouping and organizing were effective in serial learning.

A later study of McGamble and Wilson comes close to the present problem, being directly concerned with spatial localization.[7] These investigators put to their *S*s a paired-associates task with nonsense-syllables as the materials. The lists, which consisted of 18 items, were exposed in 3 horizontal rows, each syllable appearing in 1 square. Although localization was not part of the task, it exerted a pronounced effect. In the first place, localization of the items was superior to mastery of the associations. Secondly, correctly localized syllables were far better in producing recall of their partners than the others. Clearly, localization was related in an important way to association. Going in the same direction was a finding of Müller and Pilzecker that there was association of syllables with their (temporal) position in the list.[8] Most pertinent to the present investigation is a study by Gordon, who compared rates of learning of lists of syllables when they were exposed in one position, in successive positions along a horizontal line, and in regular succession around a circle.[9] The latter two conditions produced more rapid learning, but the differences were moderate.

Despite the acknowledged importance of some of the investigations just mentioned, they have not exerted a substantial effect on theory or research in recent decades. The operations of grouping and organizing, and in gen-

[5] G. E. Müller and F. Schumann, Experimentelle Beiträge zur Untersuchung vom Gedächtnisses, *Z. Psychol.*, 6, 1894, 81-190, 257-339.

[6] Stephan Witasek, Assoziation und Gestalteinprägung, *Z. Psychol.*, 79, 1918, 161-210.

[7] E. A. McGamble and Lucy Wilson, A study of spatial associations in learning and in recall, *Psycholog. Monogr.*, 22, 1916, (No. 96), 41-97.

[8] G. E. Müller and A. Pilzecker, Experimentelle Beiträge zur Lehre vom Gedächtniss, *Z. Psychol.*, 1900, Ergbd. 1, 1-288.

[9] Kate Gordon, Meaning in memory and in attention, *Psychol. Rev.*, 10, 1903, 267-283.

eral the effects of serial properties, have suffered increasing neglect. It seems appropriate to say that stimulus-response theorists hoped eventually to demonstrate that the organizing activities of learners were secondary phenomena, products of earlier associations, and that they would in time be derivable from those (allegedly more fundamental) operations.

Experiments

The present study, which takes its place with those described above, departs from them in certain respects. Unlike the earliest efforts in this area, it concentrates on the spatial properties which have been studied only in a casual way. It does not employ the method of derived lists, but relies on the analysis of certain internal aspects of serial learning, particularly errors of interference. In all of the experiments to be described, an identical list of 12 nonsense-syllables was learned. The syllables always appeared in the same temporal order. Thus the several conditions differed only with respect to their spatial properties.

General method. The syllables were projected in a dark room, one at a time, upon a translucent screen. The back of the screen was covered with a black-cardboard mask, with a pattern of windows cut in it; the syllables appeared in the windows of the screen. The patterns are described below in detail.

Each syllable was exposed for 2 sec., and the interval between exposures was 2 sec. This was also the interval between the end of one trial and the beginning of the next; there was no added pause between trials. During the intervals between exposures there was sufficient light in the room to keep the pattern of the mask visible; thus the entire pattern was always visible to *S*.

The *S* was seated approximately 3 ft. in front of the screen, which was 24½ × 24½ in., with its center at approximately eye-level. *E* read a description of the experiment, which included a detailed explanation of the order in which the syllables were to appear on the screen.

Learning was by the method of anticipation. *S* was instructed to begin anticipating as soon as he saw that the list was being repeated. In anticipating, he was to spell out aloud the next syllable before it appeared. *S* was also instructed to guess when uncertain. Responses were recorded on a prepared form. The experiment ended when the criterion of two successive correct anticipations of the entire list (from beginning to end) had been met. The syllables, in order of presentation in all the experiments here described, were: VAF, JEC, TUD, PIB, GAK, SOZ, WUJ, YEM, KAX, BIW, RUV, HEG.

Apparatus. There were two components of the apparatus: a projection-system, and a screen on which the materials were viewed.

The projection-system consisted of a rotating platform which fed the syllables into a 500-w. opaque projector (Spencer Lens Model VA Delineascope). Each syllable was so placed on the platform that it was projected onto a predetermined position on the screen. The rotary platform was powered by a synchronous motor, and so governed by a Geneva wheel that the platform rotated for 2 sec., then was

stationary for 2 sec., and so forth. A shutter-assembly was synchronized with the platform, allowing light to be projected onto the screen only when the platform was stationary. A schematic side-view of the apparatus is given in Fig. 1.

The syllables were projected onto the screen in inverse-mirror form to appear in normal orientation on the side facing *S*. In every experiment, a black-cardboard mask with a pattern of windows cut out of it was fitted to the side of the screen away from *S*. The pattern showed through sharply on the front of the screen at all times. There was a clear differentiation in the brightness of the screen, pattern, and windows. For illustration we have selected the triangular form that we studied in one of the following experiments (see Fig. 2A). The screen appeared black; the triangular pattern was yellowish (translucent wrapping paper), while the windows were white. This differentiation of brightness was present, both during exposure-periods and during the periods between exposures (due to the stray light from the projector). These brightness-relations were maintained in all of the experiments. Fig.

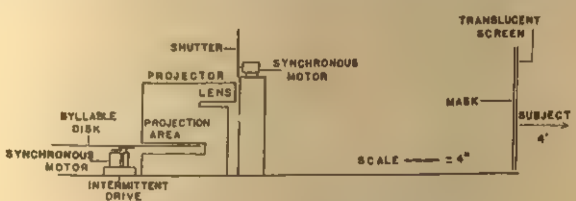


FIG. 1. SCHEMATIC VIEW OF APPARATUS

2B gives an enlarged view of the area within the dashed line of Fg. 2A, showing a single window within the contour.

The syllables were placed on the disk in the following manner. They were typed in black capital letters on white strips of paper. When the mask, with its pattern, had been fixed to the translucent screen, the syllables were moved into corresponding positions on the disk in the projection-area. Afterward the strips were cut off, leaving only small white squares of paper with the syllables printed on them, fastened to the disk by transparent adhesive tape.

For experiments I–IV, an earlier form of the apparatus was used. First, the rotating platform which fed the syllables into the optical system was turned by hand, the successive positions of the platform being determined by a notched disk. Secondly, the mask, whereby only one syllable was visible to *S* at a time, was inserted into the projector itself, in a plane parallel to and just above the nonsense-syllable platform, and held flat against the platform.

We now describe the first four experiments. Figure 3 gives a schematic sketch of the conditions.

Experiment I. Single position. The syllables appeared in succession in the same position, in a rectangular aperture, 1 in. wide and ⅝ in. high in the center of the screen (see Fig. 3A). The window always was visible to *S* as a rectangle brighter than the surround. This was the control condition, duplicating in all essentials the procedure with an ordinary memory-drum.

Experiment II: Linear array. There were 12 equidistant rectangular apertures in a horizontal line across the center of the screen. The apertures, which were ⅞ × ⅝

in. in size, were spaced at a distance of 1⅛ in. The spaces between the windows were filled with translucent tape, which also joined the ends of the line to the edges of the screen. The syllables appeared in the successive apertures, from left to right. When the twelfth syllable had been exposed at the extreme right, it was followed (without added pause) by the first syllable at the extreme left (Fig. 3B).

Experiment III. Triangular contour. The contour was an equilateral triangle, whose sides were 16 in; it was outlined by a tape 1½ in. wide. The syllables appeared in 12 apertures within the contour; they were spaced equidistantly, 4 in. from each other, and were shaped in the manner sketched in Fig. 3C. The vertex syllables occupied positions outside the contour. The positions in which the syllables appeared are indicated by the numerals in the apertures of Fig. 3; thus, the

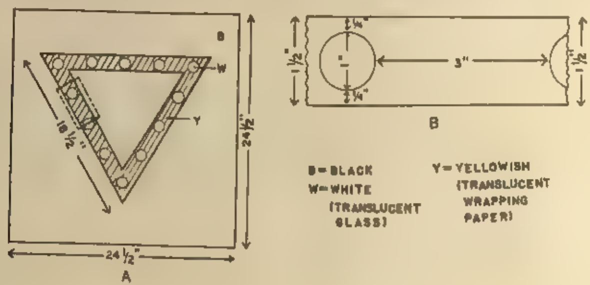


Fig. 2. Details of Presentation
Brightness-relations between screen, pattern, and window (A) and relation of window to contour (B)

first syllable, VAF, appeared in Position 1, the second syllable in Position 2, and so forth.

Experiment IV. Two positions. The syllables appeared in simple alternation in two windows, each a rectangular aperture of 1 × ⅝ in., separated by a horizontal distance of 1½ in. A strip of tape filled the space between the windows. The odd-numbered syllables appeared in the left window, and the even-numbered syllables in the right window (see Fig. 3D).

To summarize: (1) *S* learned a list of nonsense-syllables by the method of anticipation. Syllables were anticipated by spelling the letters. *S* was explicitly encouraged to anticipate whenever possible. (2) There was no added pause to mark the transition between trials. (3) The instructions contained a detailed explanation of the spatial order the syllables would follow. Thus, it was explained to *S*s of Experiment II that the syllables would appear in order from left to right, and that the last syllable would be followed, without pause, by the first member of the series in the extreme left position. In Experiment III, the *S*s were shown the position of the first syllable and were told that the series would appear in counterclockwise order. Similarly, it was explained to *S*s of Experiment IV that the syllables would alternate in position. The *S*s thus had prior information about the spatial distribution. (4) All anticipations not completely correct were scored as errors.

Swarthmore College students, both men and women, served as the *S*s. There were 10 *S*s in each experiment.

Results. The conditions differed substantially in difficulty. Those of Experiment III (involving the triangular contour) were the easiest (mean errors to criterion = 84.6), and those of Experiment IV (with two alternating positions) were the most difficult (mean errors = 236.6). The conditions of Experiments I and II were intermediate in difficulty (mean errors = 149.1 and 121.2, respectively). By the Mann-Whitney test, there were significant differences in difficulty at the 5% level between the means of Experiments I and III, II and IV, and III and IV.[10]

In Fig. 4 we have plotted the frequency of errors for the individual

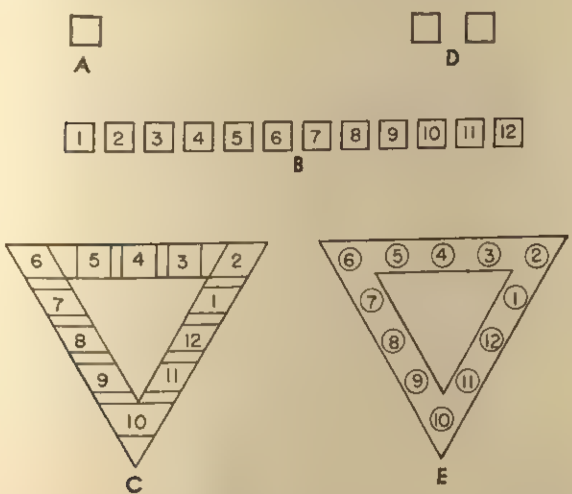


FIG. 3. SPATIAL PATTERNS OF EXPERIMENTS I–IV AND VII

(A) All syllables appear in the same aperture (Experiment I); (B) the syllables appear in 12 equidistant positions in the numbered order (Experiment II); (C) The syllables appear along a triangular contour in the numbered order (Experiment III); (D) the syllables appear alternately in two apertures, the odd syllables at the left and the even syllables at the right (Experiment IV); (E) The distribution is again triangular, but the apertures are uniformly circular (Experiment VII)

syllables in each of the experiments. The differences between the conditions tended to be consistent over the entire series.

Serial-position. Fig. 4 suggests the presence of a gradient in serial position in several of the experiments. It is more noticeable in Experiments I and II than in the others. Later we shall consider some conditions that help produce a serial gradient. At this point, we call attention to its pres-

[10] Frederick Mosteller and R. R. Bush, Selected quantitative techniques, In Gardner Lindzey (ed.), *Handbook of Social Psychology,* 1, 1954, 315-316.

ence in some experiments, despite the elimination of a temporal gap between learning-trials.

Remote errors. Variations of the spatial conditions also produced orderly differences in the frequency of various types of errors of interference.

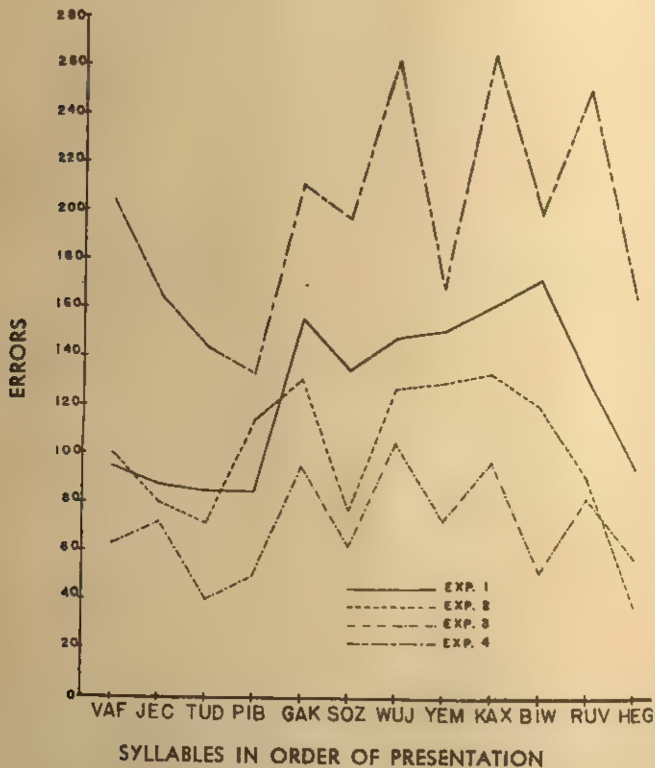


FIG. 4. FREQUENCY OF ERRORS IN RELATION TO SERIAL POSITION FOR EXPERIMENTS I–IV

Indeed, interference-errors were the most valuable source of evidence concerning the grounds of difference between the experimental conditions.

The following procedure was employed in scoring for interference- (or remote) errors. An anticipation in Position n that was identical with a syllable in Position $n + x$, was scored as an error of degree x. Similarly, if syllable $n - x$ occurred in Position n, the latter was scored as a remote error of degree $-x$.[11] We also scored for 'partial' erroneous anticipations;

[11] Under our conditions, the distinction between forward and backward errors creates a difficulty at certain points. The internal evidence suggests that certain

these consisted of two letters (in any order) common to another syllable, or the three letters of another syllable in an incorrect order. As with complete errors, these partial errors were tabulated by degree. (Once an anticipation was scored as a complete error, it was not considered as a possible partial error.)

The data appear in Table I. They consist of all interference-errors—complete and partial—grouped according to degree of remoteness. To save space we combined forward and backward errors.[12]

As described earlier, we encouraged our *S*s to respond when not certain. Fortunately for the present purpose, there were numerous erroneous anticipations. In addition, the ratio of interference-errors to the total errors was quite constant in the several experiments, ranging from 19.6-24.5%. These differences in percentage of interference-errors were not significant. We computed, for each *S*, the frequency of interference-errors and the total number of errors. By the Mann-Whitney test, none of the differences among Experiments I–IV reached the 10% level. There were, however, marked differences between the experimental conditions in the distribution of interference-errors.

First-order errors. The highest concentration of first-order errors occurred in Experiments I and II (35.4% and 37.7%, respectively, of all remote errors). They dropped in Experiment III to a level of 12.6%, and reached the level of 8.4% in Experiment IV.

Having established above that the ratio of interference-errors to total errors was constant in Experiments I-IV, we shall, in the analyses to follow, compare the frequencies of particular kinds of interference-error to the total number of *interference-* (or remote) errors. We computed for each *S* the ratio of the two sets of scores in each of the conditions under comparison, and applied the Mann-Whitney test. First-order errors (or the ratio of first-order errors to all interference-errors) were significantly lower in Experiments III and IV than in Experiments I and II (1% level); those of Experiments I and II were not significant at the 10% level.

errors, which we scored as backward, actually were anticipatory. When, for example, *S* responds to a late syllable, such as the last in the series (HEG), with an early syllable, say the second member (JEC), be probably is anticipating rather than harking back to the beginning of the series. Although it might have been more appropriate to consider this a forward rather than a backward error, we followed the latter procedure for the sake of consistency. Fortunately, this occasional ambiguity will not obscure the analysis to follow.

[12] The reader will note in Table I that interferences of a high order of remoteness occurred at times with considerable frequency. These were due almost entirely to intra-serial similarity. The interferences in question were mostly between syllables separated by even distances. Since this one-sidedness was a constant feature of all the experimental conditions, it does not impair the validity of the comparisons made above.

TABLE I

Distribution of Remote Errors (Forward and Backward, Complete and Partial) in Experiments I–IX

Experiment		Degree of remoteness 1	2	3	4	5	6	7	8	9	10	11	Total no.	% of all errors
I (Single position)	*F*	115	37	40	52	19	20	10	9	4	15	12	325	21.8
	%	35.4	11.4	12.3	16.0	5.8	6.2	3.1	2.8	1.2	4.6	1.2		
II (Linear array)	*F*	90	15	14	46	13	34	16	9	0	2	0	239	19.6
	%	37.7	6.3	5.9	19.2	5.4	14.2	6.7	3.8	0	0.8	0		
III (Triangular contour)	*F*	26	15	9	64	9	41	15	10	2	13	3	207	24.5
	%	12.6	7.2	4.3	30.9	4.3	19.8	7.2	4.8	0.9	6.2	1.4		
IV (Two positions)	*F*	46	164	25	156	11	68	13	17	3	46	0	549	22.7
	%	8.4	29.9	4.6	28.4	2.0	12.4	2.4	3.1	0.5	8.4	0		
V (Single position)	*F*	127	40	36	41	13	33	6	4	5	8	10	323	22.9
	%	39.3	12.4	11.1	12.7	4.0	10.2	1.9	1.2	1.5	2.5	3.1		
VI (Single position: Familiarization)	*F*	122	24	21	36	9	18	4	4	0	9	6	253	37.2
	%	48.2	9.5	8.3	14.2	3.6	7.1	1.6	1.6	0	3.6	2.4		
VII (Triangular contour: Familiarization)	*F*	25	16	7	32	5	22	3	17	0	0	0	127	26.0
	%	19.7	12.6	5.5	25.2	3.9	17.3	2.4	13.4	0	0	0		
VIII (Six positions)	*F*	46	12	14	53	29	76	13	4	5	10	1	263	17.2
	%	17.5	4.6	5.3	20.2	11.0	28.9	4.9	1.5	1.9	3.8	0.4		
IX (Six positions)	*F*	75	19	23	87	22	48	20	35	6	20	2	357	21.6
	%	21.0	5.3	6.4	24.4	6.2	13.4	5.6	9.8	1.7	5.6	0.6		

Second-order errors. These showed equally large differences, ranging from 6.3% (in Experiment II) to 29.9% (in Experiment IV). By the Mann-Whitney test, the ratio of second-order to total interference-errors was significantly higher in Experiment IV than in all the other conditions (1% level); none of the other differences was significant.

Odd-degree and even-degree errors. The percentage of odd-degree errors was 60.0 in Experiment I, 55.2 in Experiment II, 31.2 in Experiment III, and 17.9 in Experiment IV. By the Mann-Whitney test, all differences were significant at the 1% level, except those between Experiments I and II, and III and IV, respectively, which did not reach the 10% level.

Specific patterns of interference-errors. The incidence of errors of a given degree of remoteness is a result of interferences occurring between

TABLE II

FREQUENCY OF INTERFERENCE-ERRORS BETWEEN SELECTED MEMBERS OF THE SERIES

(Interference-errors were computed between syllables occupying given positions on the triangular contour and compared with interference-errors between the same syllables in the other experiments.)

Interference between	Experiment III		I		II		V		VI		VII	
	F	%	*F*	%	*F*	%	*F*	%	*F*	%	*F*	%
Vertex syllables	5	2.4	4	1.2	2	0.8	6	1.9	1	0.04	0	0
Vertex syllables and immediately adjacent syllables	6	2.9	70	21.5	39	16.3	75	23.2	57	22.5	6	4.3
Similar syllables	103	50.0	82	25.2	81	33.9	69	21.4	48	19.0	70	50.7

quite different members of a series. That many of the results reported above were significant despite this heterogeneity is to be attributed to the presence of strong over-all trends characteristic of each condition, but an analysis of errors between particular members of the series should provide more detailed evidence. Accordingly, we computed for each experiment the frequency of error between each pair of syllables. There were thus 132 entries (144 minus 12) for each condition. Included in each entry were forward and backward errors within each pair. The data in question tell us in detail what specific sources of interference were active in each condition, and permit us to relate these to the given figural or grouping properties.

As our base-condition we take the triangular configuration of Experiment III. (We omit comparison with Experiment IV, which presents special problems to be considered separately.) Since it was not feasible to compare exhaustively the sets of interference-errors, a more selective procedure was necessary. Accordingly we chose positions on the triangular form that were in some way outstanding. The relevant comparisons appear in Table II.

Since the vertices of the triangle form a distinct set of positions, there might be greater interference among them than among the same syllables in Experiments I and II.[13] The frequency of errors was too small to permit a conclusion concerning significance.

We reasoned that there might be lowered interference between the vertex syllables and those immediately preceding and following them. Accordingly we computed the errors between these positions. The results were clearly in this direction: both absolutely and relatively the intereferences between vertex and adjacent syllables were lowest in Experiment III. By the Mann-Whitney test, the differences between Experiments III and I, and between Experiments III and II, respectively, were significant at the 1% level; those between Experiments I and II did not reach significance at the 10% level.

A substantial proportion of the intereference-errors in all variations was due to intra-list similarity. It seemed of interest to compare the conditions in this respect. The result, which we did not anticipate, shows that the triangular condition exceeded the others, absolutely and relatively, in errors due to similarity. By the Mann-Whitney test, the ratio of errors due to similarity was higher in Experiment III than in Experiment I at the 1% level, and higher than in Experiment II at the 5% level. Experiments I and II did not differ significantly. We propose that the relative ease of the triangular task may be responsible for this result, that the lowered level of difficulty brought the learner face to face with the similarity-relations at a time when *Ss* in the other experiments were still struggling with the immediately adjacent items that produced first-order errors.

Corollary Experiments

Experiment V: Replication. Experiment V was conducted with *Ss* from Harvard and Radcliffe Colleges. In view of the shift of the locus of the *Ss* and also because of some changes in the apparatus, we undertook to see whether we could replicate the findings previously obtained. Accordingly we repeated Experiment I (the Single Position) with 10 *Ss*.

Results. An inspection of Tables I and II shows a marked consistency between the results of Experiments I and V. The data are very similar in level of difficulty (mean errors in Experiment V are 141.3), in the pattern of remote errors, and in interference-errors between selected members of the series. The similarity in the distribution of errors demonstrates that these reflect in a clear way a stable feature of the task in question.

Experiments VI and VII: Familiarization. The task of rote-learning comprises a number of sub-tasks. It requires *S* to differentiate unfamiliar items and to master

[13] If the vertex positions are outstanding, their syllables may also be mastered more quickly. It is not contradictory to assume a lower level of difficulty and a higher level of interference, provided we take relative results into account.

them in a particular order. Under the present conditions both these sub-tasks had to be mastered simultaneously. It therefore remains to be decided at what point the given spatial conditions altered the course of learning.

We were particularly concerned to determine whether the figural distributions were effective at the point of sequential ordering. This forms a major part of the task, and has been considered synonymous with the 'formation of associations' in the descriptive as well as the theoretical sense. Accordingly, we familiarized *Ss* with the members of the same series before they began the task of serial learning.

Procedure. A group of 20 *Ss* was given the procedure of familiarization, to be described below. They were then divided into two sub-groups of 10 *Ss* each. One sub-group (Experiment VI) learned the series under the single-position condition; the other (Experiment VII) under the triangle-condition. The windows on the triangular contour were now of circular shape, in contrast with the apertures of Experiment III (see Fig. 3E). We chose to compare these conditions since, as we noted above, they showed characteristic differences. We proposed to see whether the differences previously obtained between Experiments I and III would be maintained under the present conditions.

Twelve random orders of the syllables in the preceding experiments were prepared for the familiarization-procedure, each in the form of a list typed on a sheet of paper. Ten random orders of these sheets were selected, one for each *S*. *S* was instructed that the order of the syllables would vary from sheet to sheet, and that the task was to recall as many of the syllables as he could, regardless of order. He inspected each sheet for 48 sec.; thereupon he wrote down the recalled syllables for a period of 35 sec., proceeding then to the study of the next sheet, This procedure of studying the syllables and of recall was repeated until *S* had correctly recalled all 12 syllables on a given trial, or until he had 12 trials.

Serial learning followed the familiarization procedure after a lapse of 5 min. (The interval was filled with conversation unrelated to the experiment.) The *Ss*, as in Experiment V, were Harvard and Radcliffe undergraduates.

Results. The two groups were fairly evenly matched on the familiarization-task. All *Ss* assigned to the single-position condition had perfect recall on the last trial, as did all but two *Ss* subsequently assigned to the triangle-condition; the latter had errors of 2 and 4, respectively.

Familiarization substantially reduced the level of errors both in Experiments VI and VII. Single-position errors were reduced by 54.4% in comparison with Experiment I (mean errors = 68.0) and triangle-errors by 42.3% in comparison with Experiment III (mean errors = 48.8). The differences between the parallel variations with and without familiarization were significant at the 1% level.

After familiarization, the triangle-variation remained easier than the single-position variation, but only at the 10% level (using a two-tailed test). The pattern of remote errors was not, however, appreciably altered by the preceding familiarization (see Table I). Indeed, the distribution of errors remained highly constant. First-order errors continued to predominate in Experiment VI, while errors among similar syllables predominated in Experiment VII (see Table II). In all other respects there was a close correspondence between the experiments with and without familiarization. The proportion of odd-degree errors was 64.0 in Experiment VI and 31.5 in Experiment VII.

The procedure of familiarization thus did not eliminate differences between the

error-patterns in Experiments VI and VII. Since these differences cannot be referred to the familiarization phase, which was the same for both groups,[14] we conclude that the figural conditions produced an effect on the associative connections.[15]

Experiments VIII and IX: Factor of direction. Experiments VIII and IX deal with the role of a further perceptual condition, that of direction.

Method. The spatial distribution of the syllables in Experiments VIII and IX was as follows: The list of 12 syllables, identical with those of the earlier experiments, appeared in 6 rectangular windows arrayed in a straight horizontal line. Each aperture was 1 × ⅝ in., with a distance of 1⅛ in. between them. The first six syllables of Experiment VIII appeared in order from left to right, and Syllables 7-12 also appeared in the corresponding positions from left to right (*e.g.* 1 and 7 in the

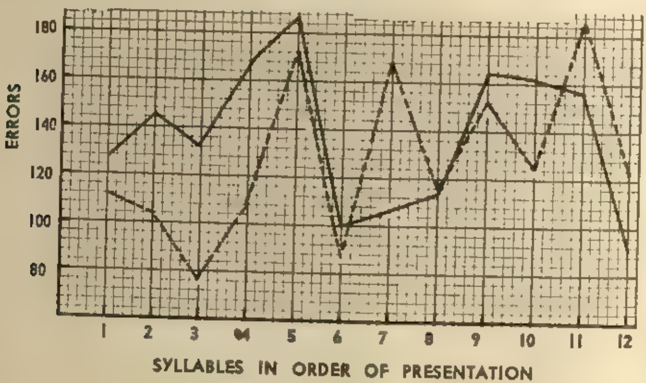


FIG. 5. FREQUENCY OF ERRORS IN RELATION TO SERIAL POSITION FOR EXPERIMENTS VIII AND IX
(Broken line, Experiment VIII; Solid line, Experiment IX.)

same window). In Experiment IX, the first half of the series appeared as above, but syllables 7 to 12 were exposed in order from right to left (6 and 7 in the same window). Thus in both experiments, two syllables appeared in each aperture.

The instructions included an explicit explanation of the spatial order in which the syllables would appear. Again, the *Ss* were encouraged to guess when they were not certain of their anticipations. The *Ss* were students at Swarthmore College, men and women.

Results. Experiments VIII and IX were of approximately equal difficulty, the mean errors to criterion being 152.9 and 165.3, respectively. (The levels of difficulty were quite similar to those of Experiments I and II.) The serial distribution of errors in the two experiments are shown in Fig. 5. Experiment IX produced a

[14] We have not studied the effect of varied perceptual conditions on familiarization itself.

[15] There was no evidence that the *Ss* ordered the syllables for themselves during the familiarization procedure. Thus the possibility of idiosyncratic orders as possible confounding factors is eliminated.

double gradient, each portion corresponding to one-half of the series, which is particularly noteworthy. The middle syllables, in Positions 6 and 7, were among the easiest, although they occupied the least favored spatial position.

Interference-errors occurred with considerable frequency—17.2 and 21.6% of all errors—in Experiments VIII and IX (see Table I). In Experiment VIII, there was a relatively low frequency of first-order errors as compared with Experiments I and II, a result probably due to the relatively higher discriminability of six positions. A more important difference emerges when we turn to errors among pairs of syllables sharing the same position (or the set of sixth-order errors). There were many more such errors proportionately than in Experiments I and II. The ratios of sixth-order errors to the total of interference-errors were, by the Mann-Whitney test (two-tailed), significantly higher in Experiment VIII than in Experiment I (1% level)

TABLE III

FREQUENCY OF INTERFERENCE-ERRORS BETWEEN SYLLABLES SHARING THE SAME POSITION

(Experiment VIII)

Experiment	Syllable-pairs						Total	% interference-errors
	1–7	2–8	3–9	4–10	5–11	6–12		
VIII (Six positions)	2	21	5	28	9	11	76	29.0
I (Single position)	0	1	1	17	0	1	20	6.2
II (Linear array)	1	1	2	29	1	0	34	14.4

and Experiment II (5% level). Table III shows the interference-errors between each pair of syllables occupying the same position. These were without exception higher in Experiment VIII than in Experiment I; with one exception the errors of Experiment VIII were higher than those of Experiment II.

If we had to rely on the evidence just described, it would be appropriate to conclude that identity of position was a source of interference. In the light of the evidence from Experiment IX, to which we now turn, it will be seen that this conclusion, although correct, is oversimplified, concealing an important point. In Experiment IX, too, each position was occupied by two syllables, one from each half of the series. The interference-errors between the pairs in identical positions are shown in Table IV. (The paired syllables were not the same as in Experiment VIII.) Syllables sharing the same position in Experiment IX did *not* interfere more with each other, either absolutely or relatively, than did the same items in Experiments I and II. By the Mann-Whitney test none of the differences between the experiments under discussion is significant at 10%.

On the basis of the data of Experiment VIII alone it would have seemed adequate to say that positional identity was responsible for the interferences obtained. In the light of the last results we must draw a revised conclusion. Both Experiments VIII and IX contained a factor of *direction*. The series of Experiment IX comprised two opposing directions in contrast with Experiment VIII, which contained two identical directions. We now see that positional identity did not produce interference when it was opposed by a difference in direction. The role of direction may not, however, be restricted to Experiment IX; it should apply with equal force to

Experiment VIII. In the latter case, organization according to direction was present to the same extent, but it coincided with identity of position. We must therefore say that the results of Experiments VIII refer to the joint effects of positional and directional identity. Added interferences occurred when both kinds of identity coincided, but not when positional identity and direction were in opposition.[16]

Experiment X: Pair-formation and proximity. We now turn to the effect produced by two conditions of proximity and the frame of reference. The syllables occupied 12 positions along a 45° line. There were six pairs of positions. The members of a pair were separated by a distance of 1 in., while the distance between successive pairs was 3 in. These measurements refer to distances along the mid-line of the diagonal strip. (The windows appeared along a 45° line to permit maximal separation between the successive pairs.)

The back of the screen was covered with black cardboard. A diagonal strip of

TABLE IV

FREQUENCY OF INTERFERENCE-ERRORS BETWEEN SYLLABLES SHARING THE SAME POSITION

(Experiment IX)

Experiment	Syllable-pairs 1–12	2–11	3–10	4–9	5–8	6–7	Total	% interference-errors
I (Single position)	4	2	2	7	4	9	28	8.0
II (Linear array)	0	2	5	8	4	7	26	10.8
IX (Six positions)	2	3	5	10	1	8	29	8.1

masking tape, 32½ × ⅞ in. was attached to the mask, extending from the lower left-hand corner to the upper right-hand corner at a 45° angle. Since the tape (while not translucent) reflected light, the entire strip was visible. Twelve parallelogram-windows, each 13/16 × 9/16 in., were cut out of the area covered by the tape. The longer sides of the window were parallel to the top and bottom of the mask, and the shorter sides coincided with the edges of the diagonal strip. Thus the windows appeared as white parallelogram-shaped apertures on a light-colored diagonal strip. The diagonal strip extended ⅞ in. beyond windows 1 and 12. The syllables appeared in successive windows from the lower left to upper right. *E* explained in advance to the 10 *Ss* (students at Swarthmore College) the order in which the svllables would appear.

Experiment XI: Pair-formation and framework. The syllables appeared in 12 *equidistant* positions along a horizontal line. They were grouped into six pairs by a frame surrounding each successive pair of syllables. The windows were 1 in. long and ⅝ in. high, and 1 in. apart. The ends of the first and last windows were,

[16] Since we have here studied only the joint effects of position and directional identity, it would be incorrect to conclude that positional identity alone is without effect, nor have we thrown light on the role of directional identity *per se*. To decide these questions it would be necessary to divorce, by experimental means, the conditions that we have studied jointly.

respectively, ½ in. from the edges of the mask. Successive pairs of syllables were enclosed by a frame that completely surrounded them. The frames were constructed of translucent masking tape which extended over the entire area included by the borders of the frame, excluding the areas of the windows. To illustrate with the first pair: above and below Windows 1 and 2 was a strip of tape ⅜ in. wide and 3½ in. long; the frame was ¼ in. wide, on the left of window 1 and on the right of Window 2. The same translucent masking tape filled the area between Windows 1 and 2. Thus, Windows 1 and 2 were presented on a common (yellow-brown) background of a brightness intermediate between the white of the windows and the opaque black of the cardboard mask. Each successive pair of syllables was enclosed by such a frame. The syllables appeared in successive windows from left to right. After the last syllable had appeared in Window 12, the sequence was repeated, the

TABLE V

RESULTS OF EXPERIMENTS X AND XI

Experiment	Mean errors	First-order errors		Second-order errors		Odd-degree errors	
		F	%	*F*	%	*F*	%
I Single position	149.1	115	35 4	37	11.4	194	60.0
II Linear array	121.2	90	37.7	15	6.3	132	55.2
V Single position	141.3	127	39.3	40	12.4	197	61.0
X (Pair-formation by proximity)	142.2	53	19.1	54	19.5	105	38.2
XI (Pair-formation by frame-inclusion)	86.9	36	21.8	41	25.7	64	38.2

first syllable appearing in Window 1 after a 2-sec. interval. There were 10 *Ss*, Harvard and Radcliffe undergraduates.

Results. The main results of Experiments X and XI, which should be considered primarily in relation to those of Experiments I, II, and V, are presented in Table V. The conditions of Experiment XI produce the fewest errors but not significantly fewer than the other conditions. There were, however, systematic differences in patterns of remote errors. (a) First-order errors were lowest—absolutely and relatively—in Experiments X and XI. By the Mann-Whitney test, the ratios of first-order errors to the total of interference-errors divide the experiments significantly (1% level) into two groups—the paired and unpaired. (b) Second-order errors occurred with the highest relative frequency in Experiments X and XI, but the differences were not uniformly significant. The paired conditions contained two sets of homologous positions, the odd and even. Homologous items, or those separated by even distances, produced higher levels of interference-error in Experiments X and XI than in comparable variations. Although these differences were moderate and not uniformly significant, they went without exception in the same direction. There was a tendency for syllables in homologous positions to interfere with each other as often, or more often, than with others that were nearer temporally and spatially. Perceptual pairing introduced a distinction between 'first' and 'second' members. We conclude that organization by spatial pattern occurred under these conditions, and that it determined the course of serial learning.

Discussion

The spatial distribution of a series of nonsense-syllables produced two kinds of effects. First, some of the conditions altered over-all difficulty. Secondly, they produced characteristic patterns of remote errors. Thus, the given perceptual conditions altered many of the effects obtained with a purely temporal series. We were able, for example, to depress the errors between temporally adjacent terms that are so typical of an exclusively temporal series, to increase second-order errors, and to alter the frequency of error due to intra-list similarity. We need now to consider how the spatial ordering of a series may function to alter the task of serial learning.

It is evident that spatial position was a source of information. When a syllable occupies a particular position, distinct from the positions of other members of the series, the learner is in possession of two sources of information at the point of anticipation—the preceding syllable (or syllables) and the position. (We are assuming, as was the case in the present experiments, that the learner knows in what position the to-be-anticipated syllable will appear.) Thus, learning under these conditions includes the formation of paired-associates between position and syllable. To account for the obtained results we may not, however, limit ourselves to associations between single items and their positions. It becomes necessary to refer to the relations between positions, or to properties of the series. This we will now attempt for intereference-errors, first with special reference to Experiments I–IV.

First-order errors occurred with highest frequency in the linear-array conditions. These were series marked predominantly by relations of sheer succession, and by the absence of other relations.[17] First-order errors were as a rule lower in series that contained further relations (Experiments III, IV and VII), a result which we have repeatedly confirmed in subsequent experiments. We conclude that first-order errors (or errors among adjacent members) occur most frequently in series whose main principle of ordering is that of uniform succession.

That errors of other degrees of remoteness were equally a function of particular series-properties, we can best illustrate with reference to the unusually high level of second-order errors in Experiment IV. This result was part of a more general effect; namely, of a high concentration of even-numbered as compared with odd-numbered errors. In this experiment,

[17] The similarity of Experiments I and II is probably to be referred to the linearity of the spatial positions. Adjacent positions (especially toward the center) may be difficult to differentiate adequately for the same general reasons that make temporally adjacent items difficult to differentiate.

even-numbered intereferences comprised 81.1% of all interference errors, a direct consequence of the division of the list into two halves.

It is reasonable to assume that the level of difficulty will also be a function of series-properties. Thus, the great difficulty of Experiment IV was doubtless connected with a division of the series into two spatial halves that cut across the linear temporal order. In general, we may say that level of difficulty will be a function of the perceived relationships among the positions. When these are clearly articulated, difficulty will decrease, while similarity of position may introduce new sources of interference.

Have we done more than demonstrate that position can function as a 'cue,' or that there was association with position? The notion of 'association with position,' although useful, also contains an ambiguity that hides a problem. Since we can point to a particular position in apparent independence of others, it might appear that association occurs with 'absolute' position. This is, of course, not the case. If we were to shift, between trials, the entire spatial array in any of our experiments somewhat to the right or upward, it is not likely that learning would be appreciably impaired, and far less likely that the pattern of errors would be modified. This should suffice to show that 'position' in this context refers largely to a relation within a spatial distribution. A given position, say on the triangular contour, has properties deriving from its relation to the contour, such as being at the vertex. It is these features of a given position that entered into association with a given syllable. The findings reported above concerning interference-errors were directly a function of these positional properties. Further, when the entire set of positions is perceived as mutually interrelated within a configuration, the associations occurring will be a function of the *Gestalt*-properties of the configuration. This is, indeed, what we find.

In Experiment III, employing the triangular form, the reduction in the difficulty of the task was not confined mainly to certain outstanding regions, such as the vertices, but rather extended to the entire series. The triangular form exerted a general effect, improving differentiation among all members by providing a set of relationships that affected the entire series. By putting before the learner a particular grouping or a particular form, we enabled him to induce a congruent kind of organization in the material to be learned. (We need not imply that all the formal properties of a visual form are represented in the learner's organization of the rote series—only further investigation can decide upon this. Rather we propose that a visual form provides the conditions for a particular organization of

the series, and that the *Gestalt*-properties of the form—or of the grouping—may be decisive.)

The factor of direction, which was studied in Experiments VIII and IX, provides additional support for the last point. Direction was a relational property that belonged to a syllable solely by virtue of its positional relation to other syllables. Since each syllable simply appeared in its position, the directional value was not a part of the syllable *per se,* but a property of the series. The results reported were a function of the relation between the directions of sub-parts of the series. They demonstrate the inadequacy of accounts that derive serial learning (in particular, associative interference) from local conditions exclusively. Finally, the grouping conditions studied in Experiments X and XI demonstrate that proximity and framework-inclusion altered in some significant respects the details of serial learning.

The presence of a serial gradient in some of the experiments, despite the elimination of pauses between trials, further demonstrates the role of properties belonging to the series. By eliminating such pauses we equalized intereefernce within the series due to temporal relations. The appearance of a gradient under these conditions cannot, it seems to us, be readily predicted from existing assumptions concerning associative interference. The learner's *knowledge* of the ends of the series constitutes one condition of the serial gradient. It seems necessary to conclude that the gradient requires, as one condition, a particular cognitive operation, namely, the identification of the beginning and end of the series and the location of items with respect to these boundaries. If this proposal is in the right direction, it follows that the gradient is not exclusively the product of associative operations; it rather presupposes a particular organization of the data to be learned. Similar findings concerning the curve of serial-position have been reported by Workman,[18] and by Wishner, Shipley, and Hurvich.[19]

Whether an outspoken serial gradient will form depends upon further properties of the series. It will tend to appear in series whose principle is that of uniform succession, whether temporal or spatio-temporal. Other properties of series will have diverse effects. Thus Experiment IV, by introducing a particular spatial organization, prevented the formation of a usual

[18] W. G. Workman, An experimental investigation of cognitive factors as contrasted with noncognitive factors in rote serial learning, Unpublished Doctoral dissertation, University of Chicago, 1951.

[19] Julius Wishner, T. E. Shipley, and M. S. Hurvich, The serial-position curve as a function of organization, this JOURNAL, 70, 1957, 258-262.

gradient, while Experiment III introduced a set of relationships unfavorable to a clear serial gradient,[20] and Experiment IX produced a double gradient.

That spatial grouping and form exerted forceful effects on the course of serial learning is all the more noteworthy when we consider that the relations bewteen positions and syllables were wholly arbitrary; nothing in the nature of these intractable materials called for these particular juxtapositions. The results testify to the fundamental importance of the spatial dimension in learning, and suggest that an exclusive emphasis on temporal condtions may be one-sided. They also raise the question of whether the functioning of the spatial dimension may not in some respects be unique. It is readily possible, for example, to transpose the structure of the present experiments into another dimension. Thus one may introduce chromatic differences within a list of syllables to correspond to the spatial differentiations. We suspect that the spatial conditions may prove by far the more powerful, although investigation must decide.

What is the specific difference between a temporal and a spatio-temporal series? A temporal series consists of successive data; a spatio-temporal series simultaneously introduces given properties with important relational characteritics. The latter enrich the structure of the series and, as we have seen, they exert strong as well as orderly effects. To the extent that such series alter the course of learning, they demonstrate the presence of operations not derivable from dyadic associations alone.

Working with simple forms and groupings, we were able to relate the data of serial learning to the properties of the stimulus. It is customary in associative studies to define the stimulus-condtions without inquiring into their content. Their perceptual and cognitive properties are regarded as 'givens' which may safely be ignored as soon as one turns to the problems of association. It could be readily shown that in practice investigators do take such properties into account; indeed to do so is a necessity. It would seem advisable to do so explicitly, and not to be content with the generality that what is perceived also can be associated

Summary

(1) We transformed the temporal series of rote-learning into a spatio-temporal series by exposing the members of a list of nonsense-syllables along different spatial contours. With this procedure we could follow the learning of the same task under different conditions of spatial distribution.

[20] A further contributing factor was the comparatively high level of error for similar syllables in Experiment III.

An identical list of 12 syllables was learned by comparable groups when the syllables appeared in a linear array of 12 positions, along a triangular contour, and alternately in two positions. The syllables also were exposed in six successive positions, the second half of the list occupying the same positions as the first half; in one variation the direction of the two halves was constant, in another the direction was reversed. Finally, the list was ordered in pairs along a straight line of 12 positions; the pairs were formed once by distinctive proximity, and once by including equidistant items in a framework. For purposes of comparison, the list also was presented in one position, in accordance with the usual procedure of rote learning.

(2) These figural conditions produced differences in the rate of learning, and they controlled in detail the patterns of remote errors. The effects were related to the properties of the respective configurations.

(3) Familiarizing the learner with the series prior to the sequential task did not alter the patterns of interference, demonstrating that the figural conditions produced the effects described at the point of sequential learning proper.

(4) Serial-position gradients appeared in the absence of added intervals between trials, a result not predictable from associative operations. This finding suggests that the discrimination of the boundaries of a series constitutes one condition of the formation of serial gradients.

(5) The spatial properties of the series altered many of the typical effects obtained with purely temporal series. The results demonstrate that the contribution of relational properties is in some respects as effective as sheer temporal order, and that interpretations of serial learning exclusively in terms of dyadic associations fails to account for the obtained findings.

(6) The results demonstrate the fundamental importance of spatial conditions in serial learning. The need to relate the study of serial learning to the properties of the stimulus was discussed.

OVERESTIMATION IN SIZE-CONSTANCY JUDGMENTS

By V. R. CARLSON, National Institute of Mental Health

The present study stems from the view that, under representatively natural conditions for the perception of depth, the perceived size of objects is invariant with distance. This proposition is based upon the assumption that size-constancy is the most likely result of biological adaptation to the natural environment. Available data do not seem to support this view, however, except in the sense that, on the average, size-constancy appears to hold as an approximate general rule. In particular, when *O* is asked to judge actual physical size ('objective' instructions), perceived size tends to increase with distance.[1] When *O* is asked to judge visual projective size ('analytic' instructions), perceived size decreases with increasing distance.[2] When *O* is asked to base his judgment upon a naïve, natural impression of size ('look' instructions), size-constancy is nearly perfect, at least on the average, with some tendency toward underconstancy.[3]

There are two related issues involved: first, whether the several kinds of instruction define variations of the same task or define different psychological tasks for *O*, and, secondly, whether cognitive attitude or functional, discriminative capacity is being measured. A conceptually idealistic situation for measuring size-constancy in terms of discriminative capacity might be stated as one in which *O* performs a functional task, under representatively natural conditions, which depends on the discrimination of sizes at varying distances, but does not call for a report (either verbal or manipulatory) reflecting the subjective appearance of size. Ideally, *O* would not be verbally aware that the task measures size-discrimination at all, and (hypothetically) his performance would indicate no systematic deviation of perceived size from actual size. When *O* is given the task of explicitly judging size, however, any response-set, attitude, or belief he may have about size-distance relationships may be invoked to a greater or lesser degree and may constitute the vehicle by means of which intellectual and motivational reactions to the situation affect his response.

Probably the major attitude that *O* is likely to manifest is the concept of per-

* Received for publication May 22, 1959.

[1] E. L. Chalmers, Jr., Monocular and binocular cues in the perception of size and distance, this JOURNAL, 65, 1952, 415-423; A. S. Gilinsky, The effect of attitude upon the perception of size, this JOURNAL, 68, 173-192; Noël Jenkin, Effects of varied distance on short-range size-judgments, *J. exp. Psychol.*, 54, 1957, 327-331; W. M. Smith, A methodological study of size-distance perception, *J. Psychol.*, 35, 1953, 143-153.

[2] Gilinsky, *op. cit.*, 173-192.

[3] Egon Brunswik, *Perception and the Representative Design of Psychological Experiments*, 1956, 67-69; J. L. Singer, Personal and environmental determinants of perception in a size constancy experiment, *J. exp. Psychol.*, 43, 1952, 420-427.

spective. In *O*'s terms, this is the notion that apparent size becomes smaller at increasing distances. That 'seeing in perspective' may be considered an attitude has been cogently argued by Gibson.[4] The present argument is that diminution in apparent size with increasing distance is a generalized cognitive attitude which we use in communicating with each other about distance. It is a culturally-agreed-upon answer to the question, "Why does an object look farther away when it is at a greater distance?" Greater distance is usually, at least since the Renaissance, represented by smaller object-size in drawings, paintings, and other pictorial media. This attitude does, to be sure, have a counterpart in natural, direct visual experience. One can notice that the image of a nearby, smaller object can cover the image of a distant, larger object; but this may be considered an exercise in sighting in diverging directions from a common point.[5] The perspective attitude need only be consistent with this experience, not necessarily a direct result of it; and, in general, visual experience in the primitive natural environment need only fail to contradict the perspective attitude in order for it to be maintained. One might even argue that the perspective attitude could not arise unless there were perceptual size-constancy to begin with, because otherwise there would be no direct way to perceive that it is a physically larger object whose image is covered by a physically smaller object. In any event, *O* is not able to judge visual angular size with anywhere near the accuracy with which he can judge physical size[6]—a rather anomalous result if the law of the visual angle were the essential basis for 'seeing in perspective.' Furthermore, the notion that the apparent size and the apparent distance of an object are interdependent in a causal way, which serves to explain over- or underestimation of size as due to over- or underestimation of distance, is by no means unequivocal.[7]

Postulating the perspective attitude, then, as *O*'s conceptual equipment for understanding instructions, both over- and underestimation in the usual size-constancy situation may be understood as attitudinal manifestations which are reasonably consistent with the instructions and the stimulus-situation, but not exclusively determined by either. First, it may be argued that objective instructions define the task as making a judgment according to the perspective attitude. From *O*'s point of view, a near object must 'look' larger than a far object for the two to be equal in physical size. The amount by which it must appear larger is postulated to be a subjective criterion. For many *O*s, it may merely mean that it is sufficient to make a match in which the near object appears not-smaller, resulting only in a slight bias in the other direction, but relative to a given experimental situation the discrepancy may be highly reliable. That is to say, the particular physical distance employed does not determine the amount of discrepancy in size required to meet *O*'s criterion, except that a greater perceived distance should in a general way require a greater discrepancy. Given several discriminably different distances in the same setting,

---

[4] J. J. Gibson, The visual field and the visual world: A reply to Professor Boring, *Psychol. Rev.*, 59, 1952, 149-151.

[5] R. B. Joynson, The problem of size and distance, *Quart. J. exp. Psychol.*, 1, 1949, 119-135.

[6] Brunswik, Distal focussing of perception: size-constancy in a representative sample of situations, *Psychol. Monogr.*, 56, 1944, (No. 254), 23.

[7] H. E. Gruber, The relation of perceived size to perceived distance, this JOURNAL, 67, 1954, 411-426; F. P. Kilpatrick and W. H. Ittelson, The size-distance invariance hypothesis, *Psychol. Rev.*, 60, 1953, 223-231.

amount of overestimation may be a fairly precise function of distance, but only because trials at different distances are not really independent, and *O* can judge the distances relative to each other.

A variety of so-called 'look' instructions have been used, calling in principle for a judgment based upon size as it appears, uninfluenced by any other knowledge about actual size or about size-distance relationships. The hypothesis presented here is that, ideally, such instructions tend to induce *O* to relinquish his subjective criterion that the near object appear larger or not-smaller, with no resulting bias in response-set. From *O*'s point of view, the nearer object need not appear large enough to be the same physical size as the far object, and his response is more completely determined by the immediate stimulus-complex. There is the difficulty, however, that instructions often have suggested something other than equality of size; namely, some kind of compensation for distance or a judgment of objects as if they were at the same distance. In terms of the perspective attitude, this implies a subjective criterion in which the nearer object must not be larger than the farther object, resulting in a bias toward underconstancy.

Primarily, then, the present experiment is an attempt to determine whether the difference between overestimation and a veridical size-judgment may be understood as a shift in attitude on *O*'s part. Consequently, *O*s were allowed to differentiate 'apparent visual size' from objective size but were not required to do so. The suppositions were: (1) If *O*'s attitude is shifted away from making a judgment in terms of objective size without introducing the implication of an analytic task, a veridical judgment with respect to physical size will result; (2) since the determinants are hypothesized to be different, the two kinds of judgment will be statistically independent in spite of structuring the situation in such a way that no difference need occur.

METHOD

Observers. The *O*s were 44 college students, who were paid for serving in the experiment, and 18 normal control patients, volunteers of the same age-range. These two populations did not differ significantly in the performances investigated here. The total group, 39 men and 23 women, was divided into four groups, each with both populations approximately equally represented, and equally divided with respect to sex. Except for these balancing requirements, assignment of individuals to groups was random.

Procedure. Using free binocular regard, *O* adjusted a near, variable triangle to match a far, standard triangle in size under two different instructions. The first group of *O*s received an instruction to match for apparent size on Days 1 and 2, and an instruction to match for objective size on Days 3 and 4. The second group of *O*s received the instruction for objective size on Days 1 and 2, and the instruction for apparent size on Days 3 and 4. At least several days, usually one week, intervened between successive testing-sessions.

For Groups 1 and 2, the angular separation between the standard and variable

triangles was always 20°. Since it might be argued that *O* would make a distinction between apparent size and objective size only as long as he could simultaneously compare the images of the standard and variable stimulus-objects, the experiment was repeated for Groups 3 and 4, treated exactly the same as Groups 1 and 2 but with an angular separation of 90°. After completing this experiment, the following additional judgments were obtained from Groups 3 and 4 on subsequent days: (a) The same apparent size and objective size judgments with the 20° separation between standard and variable; (b) judgments of projective size in both the 20° and 90° conditions; and (c) distance-judgments.

Instructions. The instructions were designed on the basis of preliminary testing to meet both of the following requirements as closely as possible: (a) if there is no natural tendency for *O* to differentiate conceptually between apparent size and physical size, he must not be forced to do so artificially; (b) if *O* does make a distinction between these two concepts, he must not be left in doubt as to which one he should apply in making his response. For 'apparent' size *O* was told:

> So adjust the variable that it looks equal to the standard in apparent visual size. It may also be equal in actual physical size at that point, or it may not—we are not concerned about that. Try to adjust the variable that it appears equal to you visually, whether you think it is equal in actual size or not.

For 'objective' size *O* was told:

> So adjust the variable that it is, as best you can judge, equal in actual physical size to the standard, so that if you were to measure them both with a ruler they would measure the same size. They may also look equal to you in apparent visual size, or they may not—we are not concerned about that. Try to adjust the variable so that you think it is the same actual size as the standard whether it appears equal to you visually or not.

With either instruction, *O* was told that the standard stimulus-object would be altered in size from trial to trial at random, that the variable would always be set for him at a randomly determined starting point, and that he could move the variable back and forth until he was reasonably satisfied with the match. He was encouraged to take sufficient time to make each setting carefully but not to worry unduly over it.

For projective size (Groups 3 and 4), *O* was instructed that the subtended angle of the variable triangle should be equal to the angle subtended by the standard triangle. The geometry of the problem was illustrated by pencil sketches, and as much explanation as necessary was given to ensure that *O* clearly understood the result he was to try to achieve by his setting of the variable.

For the distance-judgments (Groups 3 and 4), *O* was told that he was to indicate where the midway point between himself and the apparatus of the standard triangle 'seemed' to be. The triangle itself was not visible. *E* then held a 2-ft. × 2-in. strip of white cardboard vertically and said:

> Tell me when this piece of cardboard looks like it is at the halfway point. [*E* moved away from *O* toward the standard apparatus and back and forth as directed by *O*. Following this judgment *E* said], Now I wish you to indicate where you think the actual midway point might be in terms of actual physical distance. [If *O* said that where he thought the midpoint actually was and where it 'seemed' to be were the same (and most *O*s insisted that this was the case), *E* said], Well, suppose I were to tell you that this point you judged to appear to be halfway is

not where the midpoint actually is. Which way do you think I should move in order to get closer to where the actual midpoint is? [Starting from O's first designation of a midpoint, E thus forced a second judgment which was different from the first, but O determined the direction and amount of the difference.]

Stimulus-conditions and apparatus. The stimuli were two isosceles triangles cut from white cardboard with an altitude-to-base ratio of 1.39. Each triangle was mounted on an apparatus which allowed the triangle to be adjusted up and down through a slot in such a way that its visible portion could be varied continuously in size. O adjusted the variable by means of a control-knob mechanically coupled to the apparatus. Its limit of travel was 235 mm. The standard was set at three primary values, 44 mm. (76 mm. for Groups 3 and 4), 117 mm., and 159 mm. in altitude. In addition, four values were included on either side of 117 mm., spaced in steps of 3.2 mm., partly to maintain the appearance of a relatively large number of standard stimuli, and partly to obtain a measure of sensitivity to relatively small differences in the standard stimulus. In a single session, each value of the standard was presented twice in randomized order, and altogether O made 22 settings in a session, requiring 10-15 min. to complete.

The immediate background of the triangle on each apparatus was black felt, 30 in. wide and 36 in. high, with the triangle located approximately in the middle. Each apparatus was so placed on a table that the triangles were at O's eye-level. The variable triangle was 10 ft. away from O, and the standard 40 ft. away. The area was well-lighted by overhead fluorescent fixtures, but the triangles themselves were partially shielded from direct illumination. Although the triangles were distinctly and equally visible, inhomogeneities in the felt background were not detectable. The room itself was similar to a 10-ft. wide corridor but with miscellaneous objects in it. In general, cues to distance were very good.

Subjective reports. At the end of the fourth session, O was asked whether he tried to do anything differently under the two different instructions. He was also asked whether the triangles looked equal visually when he matched for objective size and whether either way of matching for size was any easier for him than the other. The primary intent of the questioning was to determine whether O thought he should compensate for distance in estimating size and, if so, in what direction, but this question was not put directly.

Additional tests. The following tests were also administered: the Otis Self-Administering Test of Mental Ability (Higher Examination); the Raven Progressive Matrices (1938, Revised Order 1956); and a body-sway test of suggestibility. The two intelligence-tests were given on different days, and the body-sway test was given twice, on different days. For the body-sway test, O was connected at the waist to an apparatus which recorded postural sway continuously. During the test-period, lasting 2 min., a tape-recorded voice suggested that O was falling forward. Control-periods with no suggestion, each lasting 45 sec., preceded and followed the test-period. During the entire session, O stood with eyes closed and heels together. Two measures were obtained, the mean height of the test-curve above the control-value and the highest point of forward sway. These tests were included on the hypothesis that, if differences in the size-constancy judgments were essentially a matter of cognitive attitude and motivational reaction to the situation, then individual differences would be correlated with performance on psychological tests of this sort.

It is not implied, however, that these particular tests are necessarily the best possible ones for the purpose.

Results

The data are expressed as the ratio of *O*'s setting of the near triangle to the given size of the far triangle in physical size, and a value greater than 1.00 means that the size of the far triangle was overestimated. Significance-tests were performed by computing the *t*-value for the difference between correlated means. The ratios averaged over *O*s for the 20° separation between standard and variable (Groups 1 and 2) are presented in Table I. The means for first and second days under the same instruction

TABLE I

Averages and Reliabilities for Ratio of *O*'s Setting of Near Variable to Size of Far Standard

(Variable-standard separation = 20°)

Instruction		Apparent size			Objective size		
Size of standard (in mm.)		44	117	159	44	117	159
Apparent size first ($N_1=16$)	Mean	0.91	0.97	0.94	1.28	1.19	1.15
	σ	0.30	0.21	0.20	0.37	0.17	0.12
	r_{II}*	0.90	0.88	0.81	0.87	0.66	0.81
	r_p†	0.84	0.84	0.73	0.85	0.54	0.75
Objective size first ($N_2=16$)	Mean	0.90	1.03	1.03	1.22	1.25	1.20
	σ	0.25	0.25	0.19	0.28	0.18	0.14
	r_{II}*	0.91	0.88	0.85	0.84	0.85	0.78
	r_p†	0.79	0.78	0.77	0.68	0.74	0.67

* Reliability coefficient between values obtained on different days.
† r_{II} with values for alternative instruction partialed out; $p=0.01$ at $r_p=0.64$; $p=0.05$ at $r_p=0.51$.

were not significantly different and have been averaged together in the table. There are minor differences due to differences in the size of the standard and between the two groups of *O*s. The 44-mm. standard tended to be underestimated relative to the other two sizes, probably because it was obviously small in contrast to the rest of the series. Occasionally, under the instruction for objective size, *O* reached the upper limit of travel of the variable when judging the 159-mm. standard, for which reason the means for this condition might otherwise have been slightly higher. The group, which received the objective-size instruction first, tended to overestimate somewhat relative to the other, but this difference is not statistically significant. The important results with respect to the average values are (a) none of the means for the apparent-size instruction are significantly

different from 1.00, and (b) the means for the objective-size instruction are all significantly greater than those for the apparent-size instruction ($p < 0.001$).

The average within-*Os* correlations between settings and the slightly differing standard sizes on either side of 117 mm. were significant, but not different for the two instructions or for the two groups of *Os*. Within-*Os* variabilities for the three primary standard sizes similarly showed little difference between instructions or between groups. Thus, there was no evidence that *O* was any more or less sensitive to differences in the standard, or that he paid more or less attention to the standard, under one condition as compared to the other.

The correlations between Day 1 and Day 2 show that *O* gave consistent and reliable judgments within each instruction (Table I). The values for apparent size are somewhat higher than those for objective size, but the scattergrams suggest that the apparent-size instruction distributed individual differences somewhat more evenly. The reliability-coefficients are statistically significant when the values obtained under the alternative instruction are partialed out. Even though performances under the two instructions are correlated with each other (average $r = 0.62$, $p = 0.01$), a significant amount of variation under each instruction is independent of variation in the other.

For the correlations with intelligence and suggestibility-scores, Groups 1 and 2 were combined and the size-ratios within instructions were averaged together over the three sizes. Since the two intelligence-tests gave essentially the same results, they were averaged together. The correlation between apparent-size estimation and intelligence was -0.44 ($df. = 30$, $p = 0.01$). With objective-size estimation partialed out, this coefficient was -0.54 ($df. = 29$, $p < 0.01$). The negative sign means that *Os* tending to score higher on the intelligence-tests tended to make their apparent-size judgments more in the direction of an underestimation of size. Neither the zero-order nor the corresponding partial correlation between objective-size estimation and intelligence was significant.

The two body-sway measures, average sway and highest point of sway, correlated highly with each other ($r = 0.95$) and were about equally reliable (0.66 and 0.73, respectively). Thus, there seems to be no advantage in using a measure of average sway, which is more difficult to compute, and only the highest point of sway was used in the correlations with the size-constancy scores. The situation tended to be just the reverse of that with intelligence. The correlation between body-sway and objective-size estimation was 0.30 ($p < 0.10$); with apparent-size judgment partialed

out, 0.36 ($p < 0.05$). The relationship with apparent-size estimation was not significant.

In general, the results for Groups 3 and 4 with the 90° separation between variable and standard (Table II) are similar to those already presented for the 20° separation, except that the group receiving apparent-size instructions first significantly overestimated the standard when matching for apparent size ($p < 0.01$) and showed a doubtful degree of independence between judgments of apparent and objective size. Their objective-size means are significantly higher than their apparent-size means

TABLE II

AVERAGES AND RELIABILITIES FOR RATIO OF *O*'S SETTING OF NEAR VARIABLE TO SIZE OF FAR STANDARD

(Variable-standard separation = 90°)

Instruction		Apparent size			Objective size		
Size of standard (in mm.)		76	117	159	76	117	159
Apparent size first ($N_3 = 15$)	Mean	1.15	1.19	1.17	1.27	1.27	1.26
	σ	0.18	0.16	0.14	0.17	0.16	0.12
	r_{11}*	0.78	0.92	0.70	0.54	0.62	0.65
	r_p†	0.71	0.90	0.58	0.41	0.48	0.33
Objective size first ($N_4 = 15$)	Mean	1.01	1.07	1.05	1.27	1.27	1.24
	σ	0.17	0.18	0.19	0.26	0.23	0.17
	r_{11}*	0.78	0.85	0.89	0.94	0.96	0.88
	r_p†	0.80	0.89	0.92	0.94	0.96	0.88

* Reliability coefficient between values obtained on different days.
† r_{11} with values for alternative instruction partialed out; $p = 0.01$ at $r_p = 0.66$; $p = 0.05$ at $r_p = 0.53$.

($p < 0.02$), however, and the apparent size means for the group receiving objective-size instructions first are not significantly different from 1.00. The difference between these two groups of *O*s would seem to represent a real interaction between order of receiving instructions and angular separation of variable and standard, since the results (not shown) for both groups under the subsequently-presented 20° condition were quite comparable to the results of the previous groups under that condition.

Projective size was overestimated, as might be expected, and moreso with the 90° separation: the mean size-ratios were 0.59 and 0.65 for 20° and 90°, respectively, as compared to a value of 0.25 for a geometrically correct projective match. The half-distance of the standard was also generally overestimated, by about 1-3 ft. Of the 30 *O*s, 23 increased further their estimates when told that their first estimates were not at the actual halfway point, and the average difference over all *O*s (0.9 ft.) was sig-

nificant ($p < 0.02$). The main purpose in obtaining the judgments of projective size and distance, however, was to ascertain their association, if any, with the apparent-size and objective-size judgments. These correlations (Table III) are based on the average of the three size-ratios for each instruction for each *O*. The pattern of relationship between the size-judgments and the distance-judgments is straight-forward: projective size tends to be correlated negatively, objective size positively, and apparent size not at all, with the distance-judgments. The only significant correlations within the size-judgments are those involving apparent size for the group of *O*s

TABLE III

INTERCORRELATIONS AMONG THE VARIOUS KINDS OF JUDGMENT
(N=30)*

	Variable-standard separation	Apparent size	Objective size	Apparent distance	"Actual" distance‡
Apparent size	20°	—	0.31	−0.02	0.13
	90°	—	0.64† 0.06	−0.08	−0.11
Objective size	20°	—	—	0.46	0.28
	90°	—	—	0.45	0.21
Projective size	20°	0.23	0.00	−0.21	−0.35
	90°	0.50† 0.16	0.06	−0.28	−0.36

* Groups 3 and 4 combined; $p=0.01$ at $r=0.46$; $p=0.05$ at $r=0.36$.

† Computed separately for group receiving apparent-size instruction first (upper value) and group receiving objective-size instruction first (lower value); N=15; $p=0.01$ at $r=0.64$; $p=0.05$ at $r=0.51$. All other values were negligibly different for the two groups.

‡ The two sets of distance-judgments were obtained differently, but the distinction between them is questionable. See text.

receiving the apparent-size instruction first. Paradoxically, apparent size for this group is correlated with both objective size and with projective size, which are not correlated with each other.

For the latter two groups of *O*s, the intelligence-test scores were, again, correlated negatively with the apparent-size judgments, but the values were not statistically significant. The body-sway suggestibility scores, however, were correlated to about the same extent as before with the objective-size judgments ($r = 0.38$, $p < 0.05$). Neither intelligence nor suggestibility was correlated with projective size.

Of the total 62 *O*s, all but 10 were reasonably clear in expressing the belief that the far triangle must appear smaller than its actual size because it was farther away. They indicated that in making an objective match the

near triangle looked somewhat larger than the far triangle in apparent visual size, but they did not first make a setting according to apparent size and then change it in order to achieve an actual size-match (or *vice-versa*). It must be remembered that these comments represent an after-the-fact appraisal on *O*'s part of what he thought he had done in the situation. Usually the statements were phrased in such a way as to indicate uncertainty. If *O* did make a verbal distinction between apparent size and objective size, however, it was always in the direction of the more distant triangle appearing smaller (or the nearer triangle larger) and, with four exceptions, he also produced an average objective-size judgment which was at least slightly larger than his average apparent-size judgment. The converse was not true. Several of the *O*s failing to profess a distinction between the two notions of size nevertheless made objective-size matches which were appreciably larger than their apparent-size matches. Altogether, seven *O*s produced average apparent-size matches which were larger than their objective-size matches. These occurred with the 90° separation between standard and variable, and the same *O*s conformed to the reverse, general trend with the 20° separation. Most of the *O*s tested both in the 20° and the 90° condition rated the latter as somewhat more difficult. None thought the 20° condition more difficult, and all expressed a judgment of approximate equality in comparing apparent size and objective size with respect to ease or difficulty.

It was necessary to force a judgment of 'actual' distance which was any different from a judgment of 'apparent' distance. After making the first estimate of where the halfway point 'seemed to be,' most of the *O*s immediately protested that their judgment of where this point might actually be was where they had indicated it seemed to be. Those who did not were questioned after making the second judgment, and in these cases it was obvious that *O* had simply repeated his first judgment. As with the majority of the *O*s, then, *E* started at the point of the first estimate, telling *O* the actual half-distance was not at that point and asking him to indicate where he would judge it to be. Thus, under the circumstances of this experiment, there was no indication that *O* made a conceptual differentiation between apparent distance and objective distance.

Discussion

It appears altogether clear that, perhaps with a few individual exceptions, naïve *O*s distinguish between apparent size and actual size. Jenkin

and Hyman found no correlation between judgments of projective and objective magnitudes.[8] The present data are in agreement with that result and in addition indicate a significant degree of independence between judgments of apparent size and either of the other two kinds of judgment. Furthermore, only apparent size was totally unrelated to the distance-judgments, and at the same time it was most coincident with an accurate judgment of physical size. Verbal reports indicated the general presumption that a nearer object must appear larger if it is to be the same physical size as a more distant object. These findings agree rather well with the view that, under non-impoverished stimulus-conditions, the natural, naïve perception of object-size is veridical and instructions to match for objective size introduce a bias in the direction of overestimation.

In the writer's view, the lack of a comparable conceptual distinction on *O*'s part between apparent distance and actual distance is an important datum. It further suggests that one way of *communicating* that an object looks farther away is to say that it looks smaller, but this statement is only a verbal substitute for saying it looks farther away. *O* can make an estimation of the relative proportion of the visual field occupied by an object.[9] Presumably it is this perception of proportion which makes it possible for *O* to report diminishing apparent size with increasing distance while at the same time functionally perceiving invariant size with varying distance. The determinants for both perceptions are present in the proximal stimulation, and there is no need to postulate an internal contradiction in the primitive perceptual system as suggested by Brunswik.[10] The contradiction lies rather in *O*'s cognitive interpretation of a difference in proportion of the field as a difference in apparent object-size. As long as a change in proportion is functionally equivalent to a change in distance, no behaviorally inappropriate response to objects in the environment occurs, and it makes no difference whether *O* reports a subjective change in apparent size, distance, or proportion.

The crucial question remains. If equal-sized objects at different distances in the natural environment satisfy *O*'s attitude that a more distant object is smaller in apparent size, why should they not do so in the experimental size-constancy situation with objective instructions? It seems

---

[8] Noël Jenkin and Ray Hyman, Attitude and distance estimation as variables in size-matching, this JOURNAL, 72, 1959, 68-76. These investigators also found, as in the present data, that projective- and objective- judgments of size correlated in opposite directions with distance-judgments.

[9] Joynson, *op. cit.*, 119-135.

[10] Brunswik, *op. cit.*, *Psychol. Monogr.*, 27-28.

necessary to postulate that *O* reacts to objective instructions in the way that he would if he were told to make a variable object somewhat larger than a given standard without specifying how much larger. There is no doubt that *O* could perform this task, and it is extremely unlikely that he would set the variable only one *JND* larger. The combination of objective instructions and *O*'s assumption of the perspective attitude provides a basis for defining the size-constancy judgment as such a task psychologically. One might assume that *O* has some kind of unconscious criterion based upon a 'weighted average' of his past experiences with spatial relationships,[11] which would at least set an approximate limit upon the largest discrepancy *O* would tolerate in apparent size and still call two objects equal in physical size. This proposition is very likely true, but it accounts neither for the occurrence nor for the amount of overestimation.

An additional factor is required, and it may be that overestimation is entirely a motivationally determined reaction to the experimental situation. Most *O*s express a great deal of uncertainty about the accuracy of their judgments and often ask how well they did—not in terms of visual or apparent size but rather in terms of how well they judged actual size. To most naïve *O*s, the ability to make an accurate judgment of physical size represents 'good' performance in the sense of a general value-judgment. Given a range of ambiguity as to what a judgment of equality can encompass, and given a general attitude that a corrective allowance for distance must be made in a given direction, amount of overestimation may simply be a function of the degree to which *O* is motivated to demonstrate competent performance in the situation. Singer found that experimentally produced frustration increased size-constancy judgments.[12] This result is at least susceptible to the interpretation that the *O*s were more highly motivated to perform well following failure. Furthermore, the effect was more pronounced at a greater distance, where an assumed need to correct for distance would be more obvious.

The effect of order of receiving instructions with the 90° separation between standard and variable is not inconsistent with the proposition that apparent-size instructions allow *O* to relinquish a corrective bias associated with objective instructions. Obviously, physical object-size is a much more commonly understood and agreed-upon concept than is apparent size. It is reasonable that *O* might more readily be induced to relinquish physical size as the conceptual basis for judgment after having made a set of responses according to that criterion. What is puzzling is

[11] Ittelson, The constancies in perceptual theory, *Psychol. Rev.*, 58, 1951, 290.
[12] Singer, *op. cit.*, 420-427.

that this complication was not also manifested with the 20° separation. *O*s considered the 90° condition subjectively more difficult, however, which may have tended more strongly to favor objective size as the criterion of judgment in spite of instructions specifying apparent size.

If this interpretation is correct, it may also be related to the lack of correlation between intelligence-test scores and the apparent-size judgments in the 90° condition. Size-constancy judgments under objective instructions have uniformly failed to correlate with intellectual measures.[13] If there is some correlation, it appears to be with apparent-size judgments. Thouless obtained a significant correlation in the same direction (negative) and of approximately the same magnitude as that obtained in the 20° condition of the present study.[14] The negative direction of the relationship is puzzling if on the interpretation that higher intelligence tends to go with lesser ability to perceive actual size. Locke has argued that an inverse relationship obtains in terms of biological adaptiveness, because more highly intelligent organisms have less need for direct perceptual constancy,[15] but accurate perception of physical size would seem to be more useful in the natural environment under normal circumstances, even for a highly intelligent organism. The basis for a relationship with intelligence may lie rather in *O*'s readiness to relinquish the conceptual criterion of accuracy in terms of objective size when the situation seems to call for him to do so. In any event, the correlation between intelligence and size-constancy judgments is apparently dependent upon instructions and the particular conditions of the experiment.

The present correlation with suggestibility indicates a source of individual variation which clearly need not reflect differences in perceptual capacity. Block and Block have demonstrated the usefulness of conceptualizing *O*'s response in the experimental situation as a reaction to *E* as an authority figure.[16] In their experiment, an implicit suggestion on the part of *E* was followed by some *O*s and rejected by others. A possible application to the size-constancy situation is that some *O*s may give artificially high or low constancy-values in an effort to conform with what they infer *E*'s expectations to be.[17] This possibility does not contradict the general hypothesis that an increase in motivational effort toward objective accuracy will lead to a greater bias toward overestimation. It is rather a qualification that the motivational 'effort' may also be of a defensive nature. Body-sway suggestibility has been related to neuroticism,[18] as have other personality-variables which have shown some relationship to size-constancy performance.[19] To some extent *O* is motivated to exhibit

---

[13] L. L. Thurstone, *A Factorial Study of Perception*, 1944, 70-71, 95; Noël Jenkin and S. M. Feallock, Developmental and intellectual processes in size-distance judgment, this JOURNAL, 73, 1960, 268-273.

[14] R. H. Thouless, Individual differences in phenomenal regression, *Brit. J. Psychol.*, 22, 1932, 234-235.

[15] N. M. Locke, Perception and intelligence: Their phylogenetic relation, *Psychol. Rev.*, 45, 1938, 335-345.

[16] Jeanne Block and Jack Block, An interpersonal experiment on reactions to authority, *Hum. Relat.*, 5, 1952, 91-98.

[17] The *role* of experimenter is referred to here rather than *E* as an individual.

[18] J. G. Ingham, Body-sway suggestibility and neurosis, *J. ment. Sci.*, 100, 1954, 432-441.

[19] Jenkin, Size constancy as a function of personal adjustment and disposition, *J. abnorm. soc. Psychol.*, 57, 1958, 334-338.

what he considers to be socially acceptable, desirable, or appropriate performance. Less well-adjusted individuals might be expected to be more defensively concerned to perform well, and Sanders and Pacht's three groups representing increasing psychopathology—normal, neurotic, and psychotic—exhibited an increasing tendency toward overestimation.[20] Again, young children might not manifest concern about performing well in a testing situation as much as would older ones, aging individuals moreso. There is evidence for a general trend in the direction of increasing size constancy scores over most of the age-spectrum.[21] Depending, however, upon the specific way the task is structured and, upon the particular *S*s involved, one might reasonably expect to obtain either overconstancy or underconstancy, and this is precisely what seems to have been the overall result of studies relating personality factors to size-constancy.[22] The evidence certainly is not overly convincing that variations in personality are associated with differences in the constancy of apparent size in the sense that there are corresponding differences in habitual, everyday perceptual constancy. It is just as plausible that *O*'s set or attitude toward his performance in the experimental situation can account for the association quite aside from any differences in perceptual capacity.

If this formulation is correct, however, it would be a mistake to assume either that research in perceptual constancy is irrelevant to personality theory, or that attitudinal and motivational factors are of no importance to perceptual theory. Psychophysical methodology may prove to be a very subtle and effective means of getting at certain personality variables, if predictions are based upon an analysis of the experimental situation as one in which interactions between perceptual attitudes and motivational tendencies are being more directly measured than are presumed characteristics of the individual's everyday, private, perceptual world. Theory oriented toward explication of perceptual response as a function of proximal sensory stimulation can perhaps legitimately consider as error those individual response-biases which cancel each other out in the population to which generalizations are to be made; but such theory must be concerned with response-biases which have sufficient commonality from one individual to another to appear as parameters in the population. One implication, for example, is that mathematical treatments of space perception, such as those of Gilinsky or Luneburg,[23] may partly represent the operation of commonly-held attitudes about spatial

[20] Richard Sanders and A. R. Pacht, Perceptual size constancy of known clinical groups, *J. consult. Psychol.*, 16, 1952, 440-444.

[21] Günter Pudritzki, Grössenkonstanz und Alter, *Z. Psychol.*, 159, 1956, 85-100; Thouless, *op. cit.*, 237-238; H. P. Zeigler and Herschel Leibowitz, Apparent visual size as a function of distance for children and adults, this JOURNAL, 70, 1957, 106-109.

[22] Jenkin, *J. abnorm. soc. Psychol.*, 334-338.

[23] Gilinsky, Perceived size and distance in visual space, *Psychol. Rev.*, 58, 1951, 460-482; R. K. Luneburg, *Mathematical Analysis of Binocular Vision*, 1947, 74-88, 103-104.

relationships rather than descriptions of naturally-occurring stimulus-response contingencies. These highly ingenious theoretical treatments would not in such circumstances be less valuable, but the behavior to which the concepts and constructs could be generalized would be different.

Summary

A near, variable triangle was matched with a far, standard triangle under instructions allowing, but not requiring, a differentiation between apparent visual size and objective size. On the average, apparent-size judgments were close to an accurate physical match, whereas the objective-size instruction clearly produced overestimation of the standard. The two kinds of judgment were correlated positively with each other, but there was also a significant amount of independent non-error variation. Objective- and projective-, but not apparent-size judgments were correlated with distance-judgments. The apparent-size estimations tended to be correlated negatively with intelligence, the objective-size estimations positively with body-sway suggestibility. The hypothesis is proposed that overestimation in size-constancy judgments is due to the interaction of an attitudinal bias and *O*'s motivational reaction to the experimental situation.

PERCEPTUAL SET AS AN ARTIFACT OF RECENCY

By WILLIAM EPSTEIN, University of Kansas, and
IRVIN ROCK, Yeshiva University

The concept of 'set' is invoked frequently in attempts to account for certain perceptual phenomena. In our work, we were concerned only with that type of set where it has been assumed that an expectation influences perceptual experience. The observer's (*O*'s) expectation concerning the nature of the coming perceptual event presumably facilitates the appearance of any percept which is consonant with the expectation and inhibits the appearance of any percept which is dissonant with the expectation. We were not concerned with that type of set in which intention to achieve a particular perceptual organization operates to affect the perceptual outcome.

In those situations where an expectancy-set is assumed to be operating, *O* is often first given a series of preparatory exposures which are intended to induce in him an expectation which in turn influences the outcome when the critical stimulus is presented. For example, Wertheimer describes a demonstration of stimulus-determined set in which *O* is first shown a pattern of dots such as:[1]

ab ab ab

In successive exposures the distance between dots *a* and *b* increases while that between *b* and *a* decreases. An *O* viewing this series will see

a b a b a b

as consisting of *a-b* pairs despite the fact that all dots are here equidistant from one another.[2]

No doubt *O* will develop an expection as a result of the preparatory series, but it does not necessarily follow that it is the expectation which plays the determining role in the critical exposure. The latter may be the result of the immediately preceding experiences *per se* and not of the expectation arising out of them. The failure

* Received for publication July 29, 1959.

[1] Max Wertheimer, Untersuchengen zur Lehre von der Gestalt: II. *Psychol. Forsch.*, 4, 1923, 301-350.

[2] For reviews of the phenomena of set see: J. F. Dashiell, A neglected fourth dimension to psychological research, *Psychol. Rev.*, 47, 1940, 289-305; *Fundamentals of General Psychology*, 1949, 454-456; J. J. Gibson, A critical review of the concept of set in contemporary experimental psychology, *Psychol. Bull.*, 38, 1941, 781-817; F. H. Allport, *Theories of Perception and the Concept of Structure*, 1955, 208-241.

of psychologists to consider this possibility is, no doubt, attributable to the fact that in such situations expectancy develops on the basis of the recent events. Hence, expectancy is coincidental with the recent past experiences, and their individual contributions cannot be distinguished.

In the experiments reported in this paper, an attempt was made to separate the influence of recency and frequency, on the one hand, from expectation *per se*, on the other.

Experiment I

In Experiment I, frequency and recency are put in opposition to expectancy. *O* is provided with a special set of instructions and with a specific sequence of perceptual experiences such that expectancy should lead to one perceptual organization and recency and frequency to a different one. He is told that he will be shown a series of four slides, of which three are of one figure (*e.g. A*) and one of another figure (*e.g. B*). The following series is then shown to him: *A, A, A, A/B* (where *A/B* is an ambiguous pattern containing *A* and *B*. See below). It is apparent that under these conditions (assuming that *O* is 'playing' by the 'rules of the game') the subjective probability of the occurrence of the *B*-event and the concomitant expectation grows with each additional presentation of *A*. Once the third *A*-figure has been presented, we may suppose that *O* has a strong unequivocal expectation of seeing the *B*-figure. In this situation, therefore, an expectancy-view would predict that *B* will be perceived. Both recency and frequency, on the other hand, favor *A*.

Materials. Two ambiguous figures were used. For the first we selected the well-known Schafer-Murphy ambiguous figure-ground pattern.[3] The drawing consists of an irregular line drawn vertically through a circle in such a manner that either half of the circle can be seen as the profile of a face. The two faces can also be drawn singly so that they may be presented separately in unambiguous form. Generally, the ambiguous pattern when presented only briefly will be perceived as either one or the other of the two profiles. To facilitate discussion, we will refer to the ambiguous pattern as A/B. The profile which in the usual reproduction occupies the left half of the total pattern facing to the right will be called *A*. The profile occupying the right half facing left will be called *B*.

The second stimulus-pattern was a modification of Boring's ambiguous figure: "My wife and my mother-in-law;"[4] a slightly different version of which was used by Leeper.[5] We will refer to the composite as *W/M* and to the unambiguous young

---

[3] Roy Schafer and Gardner Murphy, The role of autism in a visual figure-ground relationship, *J. exp. Psychol.*, 32, 1943, 336 ff.

[4] E. G. Boring, A new ambiguous figure, this JOURNAL, 42, 1930, 444-445.

[5] Robert Leeper, A study of a neglected portion of the field of learning—the development of sensory organization, *J. genet. Psychol.*, 46, 1935, 62.

wife as *W* and to the unambiguous mother-in-law as *M*. The use of this type of ambiguous material will lend our findings greater generality since it is not a reversible figure-ground pattern as in the Schafer-Murphy figure.

Apparatus. A Bell-Howell 2 × 2-in. projector was used, operated manually, but in such a way as to present the series of slides at a very rapid rate. The disappearance of one projected image and the appearance of the following one did not exceed 0.5 sec. A shutter placed in front of the lens served as the tachistoscope. The images were projected on a white screen in a well-illuminated room.

Instructions. The following instructions were read to each *O*.

I am interested in finding out what affects a person's ability to identify a visual figure or shape. I have here two figures. [Slides of Profiles *A* and *B* were presented to Groups 1 and 2 and described; slides *W* and *M* were presented to Groups 3 and 4 and described.] I have put together several series of slides. Each series contains four slides. Of the four slides in each series, three [Group 1] will represent Profile *B* [*A*, Group 2; *W*, Group 3; *M*, Group 4] and one of the four will represent *A* [*B*, Group 2; *M*, Group 3; *W*, Group 4]. In each series the picture which makes only one appearance has been placed in a different position. Thus, in one series, *A* [*B*, *W*, *M*] may be presented in the second position in the series of four. In another it may appear third in the series and so on. I wish to see what effect this will have on your ability to identify what you see.

The *O*s were asked to call out their responses, *A* or *B*, *W* or *M*, as quickly as possible. *E* called out the presentation-number prior to each exposure, *e.g.* "Ready? This is number two," etc.

To investigate the possibility that the outcome will be determined not by the specific shape of the *A*- or *B*-figures, but merely by a preservation of the direction in which it faces, the following control was employed with Groups 1 and 2. Each of the two groups was divided into two subgroups. For Subgroups 1a and 2a, the direction of the critical profile within the composite was the same as that direction of their most recent perceptual experience. For Subgroups 1b and 2b, the direction of the last unambiguous profile was opposite to the direction it faced as part of the composite. We may represent this experiment in the following manner, using the letters *L* and *R* to indicate the direction in which the profile is facing: *Group 1a: B(L), B(R), B(L), A/B; Group 1b: B(L), B(R), B(L), B/A; Group 2a: A(R), A(L), A(R), A/B; Group 2b: A(R), A(L), A(R), B/A; Group 3: W, W, W, W/M; Group 4: M, M, M, W/M.*

The *O*s were alternately placed in the several groups. By designing the experiment in such a way that each of the four figures was the experienced figure for one of our groups, we controlled for the possibility that our obtained outcome would reflect only a preference for a particular figure.

The exposure-times were 0.5 sec. for the unambiguous figures and 0.2 sec. for the composite. There were 20 *O*s in each of the 4 main groups. Every *O* performed individually and was interviewed at the conclusion of the experiment. Seventeen *O*s had to be excluded from the data of Groups 1 and 2 because the situation failed to induce an unambiguous expectancy or because their responses were equivocal. Twelve *O*s had to be excluded from Groups 3 and 4 for the same reasons.

Results. The results are presented in Table I. Consideration of the results for the Shafer-Murphy figures reveals no effect of direction. For

this reason we will disregard our subgroup divisions in presenting the results here.[6]

Only 3 of the 20 *Os* in Group 1 perceived the expected *A*-profile, while 11 perceived the *B*-profile, and 6 perceived both. Of the 20 *Os* in Group 2, only 6 perceived the expected *B*-profile, while 11 perceived the *A*-profile, and 3 perceived both. If we combine the data for Groups 1 and 2, excluding 'both' responses, we find that 9 *Os* (29%) perceived the expected figure, and 22 *Os* (71%) perceived the profile supported by frequency and recency. This difference in favor of recency and frequency is significant at the 1% level ($X^2 = 5.4$).

All of the *Os* gave clear verbal confirmation during the interview of the fact that an expectancy existed (except those disqualified as noted above). Very impressive in this regard was the fact that 8 of the 22 *Os* who had given non-expectancy responses revealed their expectant state by committing anticipatory errors on the third presentation.

TABLE I

EXPERIMENT I: RECENCY AND FREQUENCY VERSUS EXPECTANCY

Os	*N*	Expectancy	Frequency-recency	Both perceived
Group 1a (expecting *A*)	10	1	6	3
Group 1b (expecting *A*)	10	2	5	3
Group 2a (expecting *B*)	10	3	6	1
Group 2b (expecting *B*)	10	3	5	2
Group 3 (expecting *W*)	20	4	15	1
Group 4 (expecting *M*)	20	6	13	1
Totals	80	19 (27.6%)*	50 (72.4%)*	11

* 'Both' responses excluded.

The results for Group 3 were as follows. Only 4 *Os* perceived the expected mother (*M*) while 15 perceived the wife (*W*) and 1 reported seeing both. The results for Group 4 were comparable. Only 6 *Os* perceived the wife (*W*) which was supported by expectancy while 13 saw the mother (*M*), and 1 saw both. Combining the data for Groups 3 and 4, excluding 'both' responses, we find that 28 of the 38 *Os* (74%) perceived the drawing supported by recency and frequency, and 11

[6] The results do not preclude the possibility that with a larger *N*, the distribution for the 'same-direction' subgroup would differ from that for the 'different-direction' subgroup. For present purposes, however, it is clear that both subgroups yield the same trend against expectancy.

(26%) saw the figure demanded by expectancy. The obtained difference is significant at the 1% level ($X^2 = 8.6$).

The results thus quite decisively show that expectation as such—when separated from its usual concomitants in the set-producing situation—does not play a very significant role, if any, in determining the organization of an ambiguous pattern.

Experiment II

The purpose of Experiment II was to determine whether expectation played any role in Experiment I in deciding what was perceived when the composite was shown. A design such as this in which one factor is pitted against another can only yield information concerning the relative strength of the several variables. The result quite clearly shows the superiority of frequency-recency over expectancy. We cannot tell, however, from the results of Experiment I whether the selection of the alternative favored by expectancy about 27% of the time is based on the actual operation of expectation or is merely a reflection of the maximal potency of frequency-recency. In other words, even without an expectation for one organization, the other might not necessarily have been selected more than 72% of the time.

Procedure. With this thought in mind, Experiment I was repeated with one major change;[7] namely, expectation that the last figure would be *B* if the first several had been *A* (or vice versa) was eliminated. Hence, the only determinant present in Experiment II was frequency-recency. *A* should be perceived, when the composite is shown, because it has been seen several times immediately before; there is no obvious reason for *B* to be perceived at all. If the preference for *A* rises well above the 72% level of Experiment I, it would indicate that the expectation of seeing *B* in

[7] This was accomplished by omitting from the procedure those parts of the instructions which produced the set in the first experiment. In addition, the following modification of the preparatory series was introduced to insure the absence of any expectancy. The three consecutive presentations of the recency-figure were preceded by a single presentation of the alternative figure. Each of the four groups of *O*s in this experiment was shown one of the following four series: *A,B,B,B A/B; B,A,A,A,A/B; M,W,W,W, W/M; W,M,M,M, W/M*. The introduction of the initial figure was intended to counteract the possibility that an expectancy for the nonrecent alternative might develop on the basis of a 'negative recency effect.' This effect is a version of the 'gambler's fallacy' which leads to an expectation for the occurrence of a previously absent alternative when the alternatives are subjectively equiprobable. The *O*, who considers the appearance of the two perceptual alternatives to be equally probable, may, after reporting three consecutive occurrences of one alternative, begin to expect the other. We felt that the introduction of the nonfrequent alternative reduced the chances that such an expectancy would develop. At the conclusion of the experiment the *O*s were questioned about the existence of an expectancy. None of them reported that an expectancy had developed.

that experiment was playing an active role in the outcome by opposing the frequency-recency factor.

Results. The results of four groups of 10 *O*s each for both types of figures combined were as follows. Excluding 'both' responses, we find 28 instances (80%) favoring the frequency-recency alternative and only 7 (20%) favoring the other alternative. This distribution does not differ significantly from that of Experiment I. Apparently then, under the conditions of Experiment II, the effect of the recently experienced profiles and faces is not sufficiently great to bring about the selection of one particular alternative in the composite 100% of the time. Some randomness remains in the situation. It is this randomness, as we believe, that accounts for the score of 27% in Experiment I and not expectation.

Experiment III

Experiments I and II clearly demonstrate the ineffectualness of expectancy in determining what *O* sees in the test-situation. The design of these experiments does not, however, allow us to identify the crucial variable. The difficulty stems from the fact that both recency and frequency were pitted against expectancy. It is, therefore, possible to attribute the results to the influence of recency alone, frequency alone, or to the combined influence of both factors.

Procedure. In Experiment III, we kept frequency constant; that is, each of the two alternatives occurred with equal frequency prior to the test-event. Under this condition, recency was set in opposition to expectancy. As in the earlier experiments, expectancy was manipulated by administering a special set of instructions and by subjecting *O* to a particular sequence of perceptual events during the experimental situation.

Instructions. The *O*s were given the following instructions.

In this experiment, I am interested in investigating the ability of people to identify visual figures when the figures are presented only briefly. Let me show you the two figures we are going to work with.

[At this point the two unambiguous figures were presented and described, and the *O*s were then informed]: To acquaint you with the figures, we will begin with an easy task. In the first task the presentations will be made in a regularly alternating order. When this first task is completed, we will move on to much more difficult ones.

[The *O*s were asked to call out their identifications as quickly as possible.] Two groups of 20 *O*s participated in this experiment using the modified Boring figure. (The Schafer-Murphy figure was used in a corollary experiment with two other groups but with a different procedure as described below.) Every *O* performed individually. The experiment may be represented in the following way: Group 1, *M, W, M, W, M, W, W/M;* Group 2, *W, M, W, M, W, M, W/M*. As in the earlier experiments the unambiguous figures were each presented for 0.5 sec. and the composite for 0.2 sec.

A moment's consideration of the instructions which *O* received together with the nature of the events preceding the composite presentation will suggest that two different predictions concerning the outcome may be made. On the basis of expectancy, Group 1 should perceive the mother (*M*) when the composite is shown and Group 2 should perceive the wife (*W*). If recency is a controlling factor, the reverse should be the case. The *O*s should not see what they expect to see. The *O*s in Group 1 should perceive *W*, and those in Group 2, M.

Results. The results of this experiment are presented in Table II. Nine

TABLE II
EXPERIMENT III: EXPECTANCY VERSUS RECENCY WITH FREQUENCY HELD CONSTANT

*O*s	*N*	Expectancy	Recency	Both perceived
Group 1 (expecting *M*)	20	5	13	2
Group 2 (expecting *W*)	20	5	14	1
Group 3 (expecting *CW*)	10	2	8	—
Group 4 (expecting *CCW*)	10	2	8	—
Group 5 (expecting *A*)	20*	3	12	4
Group 6 (expecting *B*)	20	1	14	5
Totals	100	18 (20.7%)	69 (79.3%)	12

* One *O*, who gave an ambiguous response, was excluded.

*O*s were excluded because the concluding interview revealed that we had failed to induce an expectancy in these *O*s.

In Group 1 only 5 *O*s perceived the expected *M*, while 13 perceived the *W* and 2 reported both. In Group 2, again, only 5 *O*s perceived the *W*, while 14 perceived the *M* and 1 perceived both (see Table II). When the data for both groups are combined, excluding 'both' responses, the result shows only 10 *O*s (27%) perceived the face which was consonant with the present expectancy, whereas 27 (73%) attained the percept which was supported by recency. This difference is significant at better than the 1% level ($X^2 = 8.0$).

COROLLARY EXPERIMENTS

(a) Movement. To explore the generality of the effect of recency, two additional groups of 10 *O*s each were tested by ambiguous stroboscopic movement.

Procedure. We used a modified version of a stimulus-figure used by Wertheimer.[8] Group 3 was shown a series of stimulus-figures in which the movement was unambiguous, *i.e.*, a line was seen to move from a horizontal to a vertical orientation in either a clockwise (*CW*) or counterclockwise (*CCW*) direction as follows: *CW*, *CCW*, *CW*, *CCW*, *CW*, *CCW*, *CW*, *CCW*. This series was followed by a stimulus-

[8] Max Wertheimer, Experimentelle Studien über das Sehen von Bewegung, *Z. Psychol.*, 61, 1912, 218ff.

figure in which the movement was ambiguous, *i.e.*, could be seen in either direction. The *O*s of Group 4 were treated in the same manner except they were given the training series in the following order: *CCW*, *CW*, *CCW*, *CW*, *CCW*, *CW*, *CCW*, *CW*. In all other respects the procedure was the same as for Groups 1 and 2.

Results. The results are presented in Table II. In Group 3, 8 *O*s gave recency-responses, perceiving movement in the *CCW* direction. In Group 4, 8 gave recency-responses, perceiving movement in the *CW* direction. This difference in favor of recency is highly significant (X^2 for Groups 3 and 4 combined = 7.2 which is significant beyond the 1% level).

(b) Recency. A similar experiment using the Schafer-Murphy figures was performed before the technique with movement had been worked out.

Procedure. Instead of showing the preparatory figures in the sequence *A, B, A, B, A, B,* they were shown in the sequence *AAA, BBB, AAA, BBB, AAA.* Here the *A*-figure is not only more recent, but it is somewhat more frequent as well.[9] The last profile seen by each *O* faced in a direction opposite to that obtaining in the composite for the same profile.

Results. The results for two groups of 20 *O*s each are given in Table II. Eleven *O*s had to be excluded because they did not develop the requisite expectancy. In Group 5, which expected to see Profile *A*, 12 *O*s perceived Profile *B*, only 3 perceived *A* and 4 reported seeing both simultaneously. In Group 6, which was expecting to see Profile *B*, 14 *O*s reported *A*, only 1 saw *B*, and 5 saw both profiles simultaneously. Combining the data for Groups 5 and 6, excluding 'both' responses, we find that only 4 out of 30 *O*s (14%) perceived in consonance with their expectancy; and that 26 (86%) attained the percept which was supported by their immediately recent perceptual experiences and their concomitant memory traces. This difference is significant at beyond the 1/10% level ($X^2 = 14.2$). The result thus confirms that of Groups 1 and 2.

The significance of these findings for our assessment of the role of expectancy is plain. They support the first two experiments in showing that expectation as such does not seem to influence the outcome. Furthermore, the result clearly suggests the critical role of recency. We may now say that given a number of distinct, recent events (and traces thereof), it will be the most recently implanted trace which will control the perceptual organization. The reader will remember that the *W*, *CCW*, and *B* in Groups 1, 3, and 5, respectively, and the *M*, *CW* and *A* in Groups 2, 4, and 6, respectively, were only the most recent percepts, not the only recent perceptual events.

Experiment IV

We now turn to the question of frequency. In the main, we wished to know whether frequency could influence the perceptual outcome apart

[9] As the results of the next experiment show, however, the slight difference in frequency here could not have been of much importance.

from recency. We already have some information on this point. In Experiment III, 79% of the responses favored the recency-alternative; in Experiment I, 72% favored the recency plus frequency-alternative. Thus recency alone was as effective, or more so, as recency plus frequency. This suggests that the frequency with which a prior stimulus-pattern was encountered was of little consequence. In the present experiment we sought to examine this question more directly. At the same time we were interested in observing whether the influence of recency would be reduced when opposed by an inhibitory alternative which occurred frequently in the experimental situation.

This design created a situation in which each experimental group had a similar equivocal expectancy and an identical immediately recent experience. The independent variable is the frequency with which the relatively non-recent figure appeared for the several groups. We were interested in observing the effects that this might have on the strength of the recency-effect. It could very well be that the effectiveness diminishes as the alternative event occurs with increasing frequency.

Instructions. The following instructions were given to every *O*.

I am interested in finding out what affects a person's ability to identify visual figures. [*A* and *B* or *W* and *M* were presented and described.] I am going to show you a fairly long series of slides containing these two figures. The slides will appear in an arbitrary random order similar to what you would achieve if you selected the order by tossing a 'fair' coin.

For this reason you should not be surprised or misled by seeming regularities or irregularities in the series. You might see a long string of one figure, then a single alternative, then another few of the first and a long series of the second figure, etc. It is all in the way the coin falls. There are two other factors which I will vary in a random way. One is the amount of time that the slide is exposed and, secondly, the direction in which the profile will face. Sometimes the figure will be exposed for a relatively longer interval, sometimes for a very brief period. At times the profile will face left; at other times it will face right. Any questions? Ready? Try to respond as quickly as you can.

Three different degrees of frequency were pitted against the single recent presentation. The non-recent figure was shown to *O* either 3, 6, or 12 times consecutively before the figure which was immediately to precede the ambiguous test-figure. There were six groups of *O*s in this experiment. Each group performed under a different set of conditions which may be described as follows:

Group 1 (*N*–17), *A, A, A, B–A/B*
Group 2 (*N*–15), *M, M, M, W–W/M*
Group 3 (*N*–15), *A, A, A, A, A, A, B–A/B*
Group 4 (*N*–15), *M, M, M, M, M, M, W–W/M*
Group 5 (*N*–15), *A, A, A, A, A, A, A, A, A, A, A, A, B–A/B*
Group 6 (*N*–15), *M, M, M, M, M, M, M, M, M, M, M, M, W–W/M*

Every unambiguous figure was presented for 0.5 sec. The composite was pre-

sented for 0.2 sec. The direction in which the *A*-profile faced in the series presented to Groups 1, 3, and 5 varied in a random manner. In the series seen by Groups 3 and 5, the *A*-profile faced in both directions an equal number of times. For all three groups the direction of the final *B*-profile was opposite to the direction of *B* obtaining in the composite.

In this design, especially as it applies to Groups 5 and 6, one must be aware of the possibility that the *O*s after seeing a long uninterrupted series of *A*-profiles or *M* drawings may, once the sequence is interrupted by the final *B*- or *W*-figure, begin to expect a similar unbroken series of *B*- or *W*-figures. Were this the case, we would not have achieved the aims of our experiment. In the interview which concluded the experiment, we were careful to question the *O*s about this possibility.

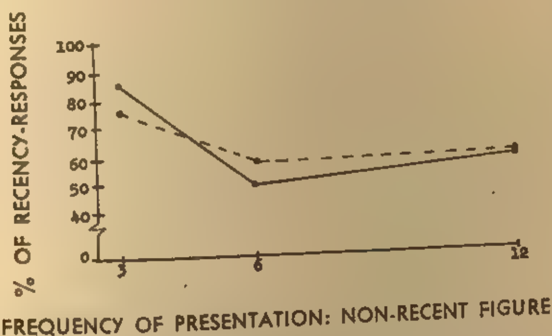


Fig. 1. Percentage of Recency-Responses when Non-Recent Figure Occurred at Varying Frequencies
——— Schafer-Murphy Profile Figures; ---- Boring's "Wife and Mother-in-Law" Figure

In none of the cases did we find evidence of an expectancy for one or the other alternative. In any event it should be noted that the results of Experiment I–III give us little reason to fear that an existing expectancy could have significantly influenced the outcomes thereby exaggerating the strength of recency.

Results. Let us first consider the results for Groups 1, 3, and 5. The response due to recency diminished in strength with the increased frequency of the relatively non-recent perceptual experience (see Fig. 1). In Group 1, 13 of the 17 *O*s gave the recency-response (perceived *B*), only 2 *O*s saw the *A*-profile which was supported by frequency, and 2 saw both. In Group 3, 6 of the 15 *O*s gave the recency-response and saw *B*, 6 reported the *A*-profile, and 3 saw both. Finally, in Group 5 only 7 of the 15 *O*s perceived *B*, while 5 attained the percept supported by frequency, and 3 saw both. The percentage of *O*s giving recency-responses, excluding 'both' responses, was 87%, 50%, 58% for Groups 1, 3, and 5, respectively.

The outcomes for Groups 2, 4, and 6 were comparable. In Group 2, 11 of the 15 *O*s saw the wife (*W*) which was favored by recency, 3 saw the mother (*M*) which was supported by frequency, and 1 saw both. In Group 4, 9 *O*s gave *W*, the recency-response; 6 gave *M*, the response favored by frequency, and none saw both. In Group 6, 9 *O*s perceived the *W*, 5 saw *M*, and one saw both. If the 'both' responses are excluded, then the percentage of *O*s giving recency-responses was 78%, 60%, and 63% for Groups 2, 4, and 6, respectively.

In both cases, then, we found a drop in recency-responses when the frequency of occurrence of the relatively non-recent figure increased from 3 to 6 presentations. A Chi-square analysis of the significance of the difference between the number of recency-responses obtained in Groups 1 and 2 combined and the number of recency-responses obtained in Groups 3 and 4 combined yielded a value of $X^2 = 5.41$ which is significant at better than the 2% level. A similar analysis applied to the differences between the distributions of Groups 1 and 3 alone and Groups 2 and 4 alone did not yield significant values of X^2. An increase in the number of occurrences of the non-recent percept to 12 did not, however, lead to a further reduction in the strength of the recency response. It seems that the influence of recency is so powerful that it cannot be mitigated too severely by the opposition of frequency. This is very striking when one keeps in mind the fact that for Groups 5 and 6 the ratio was 12:1. Despite this state of affairs, the average percentage of recency-responses for these two groups is 60.5%.

Discussion

The investigations reported above have failed to reveal any effect of expectancy as such. The design of our experiments, however, created conditions of rivalry in which expectancy competed with recency and frequency. The reader may wonder, therefore, whether expectancy would exercise an influence if it is unopposed by any other factor. The evidence presently available indicates a negative answer to this question. Leeper, using a modified version of the Boring-figure, demonstrated the ineffectiveness of expectancy induced by verbal instructions in determining which alternative is seen.[10] We have repeated Leeper's experiment with several variations in design and have included the Schafer-Murphy figures in our investigation. The results we obtained were almost identical with those reported by Leeper. Prior verbal preparation had no effect on the

[10] Leeper, *op. cit.*, 1935, 41-75.

perception of ambiguous figures. Expectancy was powerless despite the absence of any opposing variable.[11]

It may be objected that expectation under these circumstances is not sufficiently specific, *i.e.* *O* does not at the time of the test have available a sufficiently concrete representation of that which he is expected to see. For this reason a much more stringent test of the role of expectancy alone was undertaken. *O* was first shown both alternatives of either the Schafer-Murphy or the Boring-figure. Then after a 5-min. interval (which we knew eliminated the recency-effect under these conditions), he was told that he would be shown again the last figure he had seen. Despite the fact that here the expectation was quite specific and concrete and unopposed by any other factor, there was no favoring of the 'expected' alternative. Each was seen with about equal frequency.

The results of the present experiments indicate that recency is the critical variable in the 'expectancy'-effect. The nature of the most recent perceptual event has effectively controlled *O*'s experience during the test. Experiment IV has shown that increasing the frequency with which the non-recent alternative occurs results in a surprisingly slight diminution of the effect of recency. It appears from our results that the influence of recency is not greatly weakened even when opposed by high relative frequencies.

As to the generality of our findings, there are two aspects to consider. (1) The first is the question of the range of situations which may be expected to exhibit such effects. We have already shown that surprisingly similar effects are obtained with two different types of ambiguous forms and with an ambiguous-movement situation. (2) The second is the extent to which our findings can be considered applicable to all those experiments on expectancy-set which have been reported in the past. Actually, there are not too many such experiments which clearly entail perception as against interpretation, guessing, or judgmental processes. An experiment such as that of Levine, Chein, and Murphy showing that food-deprived *Ss* make a greater number of food responses to blurred pictures seen behind a ground glass screen is more a matter of interpretation than perception.[12] Here the set induced by the need determines what *S* interprets the vaguely perceived picture to be; it does not determine the

[11] It may be of interest to note that when Leeper gave his *Os* perceptual preparation with one of the unambiguous figures (expectancy absent), this preparation determined the perception of the composite almost entirely. We have repeated this experiment with comparable results.

[12] Robert Levine, Isidor Chein, and Gardner Murphy, The relation of the intensity of a need to the amount of perceptual distortion: A preliminary report, *J. Psychol.*, 13, 1942, 283-293.

shape-perception as such. In the same category are those cases in daily life often cited as instances of set in perception: walking through a dark forest, we "see" a tree as a threatening person; we incorrectly identify a distant person as a friend for whom we are expectantly waiting. Expectation obviously does operate here, but the mechanism is not understood.

Of the experiments that bear on perception, only those of Gottschaldt, Leeper, and Wertheimer, showing an effect of prior exposures on the perceived direction of stroboscopic movement similar to the experiment described above, and the demonstration by Wertheimer on grouping of dots, described earlier, are parallel to our procedures in the sense that the expectation is specific to the preparatory stimulus-objects shown.[13] Some of Gottschaldt's findings could be explained on the basis of the recency-principle, but on the whole there is reason to doubt the reliability of these findings because of shortcomings in the experimental design. Leeper's and Wertheimer's effects can certainly be attributed to recency rather than expectancy.

The remaining demonstrations of expectancy differ from our type of procedure in that the expectation is created by preparatory items which are from the same conceptual category as the ambiguous stimulus, but none of them is identical with either of its organizations. Experiments by Zangwill, Siipola, and Bruner and Minturn are of this type.[14] (Whether these experiments entail a genuine effect on *perception* is debatable.) It remains an open question whether the recency of the non-specific items can explain these effects instead of expectation as such. In any case, here too the mechanism of expectancy is not understood.

The present experiments have direct relevance to several of the present-day theories of perception. The information-theory of perception conceives "of the perceptual process as a cycle of hypothesis-information—trial and check of hypothesis—confirmation or non-confirmation."[15] Central to this analysis is the concept of an hypothesis.[16] By hypothesis these authors

[13] Kurt Gottschaldt, Ueber den Einfluss der Erfahrung auf die Wahrnehmung von Figuren: I, *Psychol. Forsch.*, 8, 1926, 261-317; II, *ibid.*, 12, 1929, 1-87; Leeper, *op.cit.*, 1935, 41-75; Wertheimer, *op.cit.*, 1912, 161-265; *op.cit.*, 1923, 301-350.

[14] O. L. Zangwill, Attitude and recognition, *Brit. J. Psychol.*, 28, 1937, 12-17; E. M. Siipola, A group study of some effects of preparatory set, *Psychol. Monogr.*, 46, 1935 (No. 210), 27-38; J. S. Bruner and A. L. Minturn, Perceptual identification and perceptual organization, *J. gen. Psychol.*, 53, 1955, 21-28.

[15] Leo Postman, Toward a general theory of cognition, in J. H. Rohrer and M. Sherif (eds.), *Social Psychology at the Crossroads*, 1951, 251.

[16] See also J. S. Bruner, Personality dynamics and the process of perceiving, in R. R. Blake and G. V. Ramsay (eds.), *Perception: An Approach to Personality*, 1951; Bruner, On perceptual readiness, *Psychol. Rev.*, 64, 1957, 123-152.

mean "in the most general sense, expectancies or predispositions of the organism which serve to select, organize, and transform the stimulus information that comes from the environment."[17] The percept which finally is attained is the outcome of a mutual accommodation of sensory stimulation and hypothesis. Thus, "if the information confirms the hypothesis, a stable perceptual organization is achieved. If the information fails to confirm the hypothesis, the hypothesis shifts until it is confirmed and a stable perceptual organization is achieved."[18] A rather similar theory is presented by the transactionalists who argue that perceptual experience depends upon the assumptions derived from past experience which are brought to bear on the stimulus.[19]

It would be improper to enter into a full-scale criticism of this thesis without giving it a more thorough presentation, but it is clear that our work points up the need for caution in its application to perceptual phenomena. We have shown unequivocally that the occurrence of a percept which is consonant with an expectancy, "specific hypothesis" or "assumption," cannot be taken as evidence supporting the belief in a causal relationship between the expectancy and the reported perceptual experience. It has been possible to find a less complicated causal relationship, thus eliminating the need for any additional intervening constructs.

Even greater grounds for caution are to be found in the fact that in many of our cases *O* attained a stable perception organization for which he was not 'tuned in.' An 'infirming' percept was attained despite the fact that the test-event took place under marginal perceiving conditions. The brief exposure-time of the composite only could have allowed for the reception of partial information. Yet we are told that, according to the theory, "the less of appropriate stimulus-information is available, the more the perceptual organization is determined by the dominant hypothesis."[20] Our findings clearly fail to support this principle. This is especially striking since, in the present studies, the expectancy or assumption was of a very specific nature. The reader will remember that *O* had an opportunity to see and hear a description of both alternatives. This is to be contrasted with the more generalized sets which are usually invoked to explain certain perceptual phenomena. It seems to us, there-

---

[17] Postman, *op.cit.*, 1951, 249.
[18] Postman, *ibid.*, 251.
[19] F. P. Kilpatrick (ed.), *Human Behavior from the Transactional Point of View*, 1952, 87-94.
[20] Postman, *op.cit.*, 1951, 260.

fore, that this analysis only serves to emphasize the need for a thorough investigation of 'set' before we presume to determine its place in perceptual theory.

Although our results can thus be considered as inconsistent with these two empiristic schools of thought, an effect of recency is, of course, an effect of past experience. Rather than interpreting perceptual phenomena in terms of broad hypotheses, assumptions and the like, however, our findings suggest that we should examine the role of specific memory-traces. Support for this position comes from the work of Wallach, O'Connell, and Neisser, who have demonstrated that specific past experience can bring about the perception of a two-dimensional pattern as three-dimensional.[21]

Before it will be possible to explain our results on the basis of the influence of specific traces, however, there are two problems which must be faced. The first is the problem of how a memory-trace can determine a perceptual outcome when ordinarily the trace is only aroused on the basis of similarity with the *organized* percept. (See Wallach and Zuckerman and Rock for a discussion of this logical difficulty.)[22] The second is the problem of why the recent trace plays such an all-powerful role.

Summary

Experimental situations were created which allowed for the separation of recency and frequency, on the one hand, from expectancy on the other. The traditional stress on expectancy as a determinant of perception was found to be misplaced. Our findings demonstrated that it was the immediately recent perceptual experience which controlled the finally attained percept. The consequences of this result for perceptual theory were considered.

---

[21] Hans Wallach, D. N. O'Connell, and Ulric Neisser, The memory effect of visual perception of three-dimensional form, *J. exp. Psychol.*, 45, 1953, 360-368.

[22] Wallach, Some considerations concerning the relation between perception and cognition, *J. Person.*, 18, 1949, 6-13; C. B. Zuckerman and Irvin Rock, A reappraisal of the roles of past experience and innate organizing processes in visual perception, *Psychol. Bull.*, 54, 1957, 269-296.

INFLUENCE OF COLOR ON THE DISCRIMINATION OF SWEETNESS

By Rose Marie Pangborn, University of California, Davis

The psychological literature contains much information regarding the affective value of colors. In a recent experiment, 40 *Os* judged the affective values of 316 different color-specimens on a hedonic scale of 11 points ranging from most unpleasant to most pleasant. Preferences were highest in the region of green to blue and lowest in the region of yellow and yellow-green, when brightness and saturation were held constant.[1] There is little or no tendency, however, for people to choose foods on the basis of color-preference alone since custom and habit combine to produce in the mind of the consumer an association between color and acceptability. Colors are identified with certain foods and processes and are acceptable or unacceptable according to the product and its intended use. In most experiments with food, color is considered an incidental attribute, less important than odor, taste, texture, size, or other aspects of the product. Although it is generally believed that a color-flavor relationship exists in foods, reports regarding the nature of the relationship are meager and few in number.

In an experiment designed to test color-associations with the flavor of food, Foster had 40 consumers evaluate two types of dry-snack products: speckled vs. unspeckled.[2] Although identical in composition, the appearance and flavor preferences for the speckled sample were significantly lower than for the unspeckled. When, however, the consumers were asked which of the two "charcoal-grilled" samples they liked the better, appearance- and flavor-preferences were higher for the speckled sample. Schutz reported that 52 *Os* preferred the appearance of orange-colored orange juice over juice of a distinctly yellow color, but taste-preferences were identical. It should be noted that both yellow and orange are within the range of color acceptability for orange juice. In the same experiment, Schutz demonstrated that the flavor-scores assigned to an inferior juice could be raised if it were artificially colored to resemble a better quality. He concluded that "spurious conclusions about food pref-

---

* Received for publication July 7, 1959. The investigation was supported in part by a grant-in-aid from the Sugar Research Foundation, Inc., New York.

[1] J. P. Guilford and P. C. Smith, A system of color-preferences, this Journal, 72, 1959, 487-502.

[2] Dean Foster, Psychological aspects of food colors from the consumer's standpoint, Mimeo. report distributed by U.S. Testing Co., Inc., Hoboken, New Jersey, 1956.

erences may be reached by considering color independently of flavor factors, and colors can be experimentally manipulated to serve as standards of good quality."[3]

Kanig, in 1955, required 200 students of pharmacy to identify variously colored and flavored syrups by taste.[4] Few of them correctly identified flavorings presented in colorless syrups, and even fewer responded correctly when solutions had an atypical color. In a study of the influence of past acquaintance upon color and taste of a product, Dunker observed that white chocolate tasted less like chocolate than did the customary brown product.[5]

The colors associated with certain flavors are not always those which occur naturally with the food. The following examples were cited by Johnson. (a) When natural carotene content is low, butter is artificially colored; (b) Mint-flavored ice cream would be white if green coloring were omitted; (c) Orange sherbet needs much more color than natural orange juice imparts to it.[6]

In the experiments described in the present paper, panels of trained and untrained *O*s evaluated paired samples of fluids varying in sugar content and flavoring. The influence of adding various coloring to these fluids on the discrimination of their flavors and sweetness was determined.

Method and Procedure

Experiment I. The solutions, presented to *O*s for comparison, consisted of various concentrations of sucrose, citric acid, imitation flavoring, and artificial coloring in freshly distilled water. Initially, imitation apricot and artificial, orange food-coloring were used. The study was then extended to include peppermint extract with green coloring, and cherry flavoring with red coloring. In each study, the objective was to determine the influence of the artificial coloring on *O*'s discrimination of flavor and sweetness.

A highly trained group of 12 *O*s (8 men and 4 women), hereafter referred to as the "trained panel," evaluated the solutions. They were served in booths which were maintained at a constant temperature and humidity. The experimental sessions were held daily at 9:30 A.M. and lasted for about 10 min. At every session the *O*s tasted 4 pairs of solutions in coded, 50-ml. beakers. They were instructed to compare the pairs and to report when the first was greater, equal to, or less than the second in sweetness and in flavor.[7] The *O*s were informed that there would be color variations in the solutions. On alternate days, overhead, daylight illumination was used and the colored solutions, which appeared bright and distinct, were easily differentiated from

---

[3] H. G. Schutz, Color in relation to food preference, *Color in Foods: A Symposium*, Ed. by K. T. Farrell, J. R. Wagner, M. S. Peterson, and G. M. Mackinnery, 1954, 21.

[4] J. L. Kanig, Mental impact of colors in food studied, *Food Field Reporter*, September 19, 1955, 23, 57.

[5] Karl Dunker, The influence of past experience upon perceptual properties, this Journal, 52, 1939, p. 263.

[6] A. H. Johnson, Significance of color in dairy products, Mimeo. report distributed by National Dairy Res. Lab., Inc., Oakdale, Long Island, New York, 1956, 13.

[7] Adaptation of a method described by R. M. Valdes, E. H. Hinreiner, and M. J. Simone, Effect of sucrose and organic acids on apparent flavor intensity: I. Aqueous solutions, *Food Technol.* 10, 1956, 282.

the colorless ones. During the other set of alternate days, red illumination so effectively masked color differences that all sets appeared uniformly red.

The stimulus-solutions, all of which contained 0.15% citric acid, were prepared according to the plan shown in Table I. The three sets of stimulus-pairs were presented as shown in Table II. Paired comparisons listed under "Duplicates" and "Double variations" were included for control purposes. The five combinations were repeated in the experimental design in a ratio of 76:38:38:19:19, respectively. The order in which the samples were presented within a pair and the order in which the four pairs were arranged within a set was randomized. Each combination was served in reverse order half of the time to compensate for any first-sample bias. To stimulate careful judgments, the *O*s were given the correct responses immediately after each trial and record of their performances was posted each day. To test the effect of training, various combinations of the design were submitted to an untrained group of *O*s.

Since the *O*s had a one-third probability of correctly guessing either the sweetness or the flavor response, the method used to determine significance was the same as that used in triangular testing.[8]

Experiment II. A group of 10 students (5 men and 5 women) between 19-24 yr. of age judged aqueous solutions and pear nectar from 4:30 to 5:30 P.M. from Monday through Thursday for three weeks. All tests were so designed that the *O*s, in groups of five, served alternately as *O* and *E*. Upon the completion of one set of tastes (6-8 min.), the *O*s and *E* would alternate, thereby reducing fatigue and monotony. The *O*s and *E* were under constant supervision to assure uniformity of conditions.

At every session, the *O*s were given four pairs of stimulus-liquid in coded, 50-ml. beakers and instructed to circle the number of the sweeter sample within each pair. Only sucrose and food coloring were added to the fluids according to the plan shown in Table III. Within each stimulus-set, uncolored samples were compared with samples containing red, green, or yellow food coloring. Uncolored pear nectar was opaque-white whereas colored nectars were pastels. All paired combinations (16) within each set were judged by every *O* at every test-period. With the exception of Set III, there was always a sucrose difference between pairs. Set III was included as a control, since identical pairs would force the *O*s to make a decision on the basis of color alone, consequently they were unaware of the presence of identical samples until the termination of the entire experiment. The main objectives of this test were: (a) to determine whether specific colors facilitated or reduced ability to discriminate between levels of sweetness; and (b) to determine the extent, if any, of the color-sweetness association. The method of Chi-square was applied to determine whether observed values differed significantly from expected values.

Results and Discussion

Experiment I. To facilitate interpretation of results, the responses of the *O*s were subdivided as shown in Tables IV, V, and VI. In Table IV,

[8] E. B. Roessler, Jean Warren, and J. F. Guymon, Significance in triangular taste tests, *Food Research*, 13, 1948, 503-505.

TABLE I

Stimulus-Solutions (A-H) and Plan of Their Preparation

Stimulus-Set	% Sucrose			% Flavoring						% Coloring			
				apricot		cherry		peppermint		yellow	red		green
	8	10	12	0.1	0.2	0.5	1.0	0.15	0.30	1.0	1.0	2.0	1.0
I	—	ACEG	BDFH	ABEF	CDGH	—	—	—	—	EGFH	EGFH	—	—
II	—	ACEG	BDFH	—	—	ABEF	CDGH	—	—	—	—	EFGH	—
III	ACEG	—	BDFH	—	—	—	—	ABEF	CDGH	EGFH			EGFH

TABLE II

STIMULUS-SOLUTIONS (A–H) AND THEIR PAIRED COMPARISON

Color variations	Flavor variations	Sucrose variations	Duplicates		Double variations
AE	AC	AB	AA	EE	AD
CG	BD	CD	BB	FF	CB
BF	EG	EF	CC	GG	EH
DH	FH	GH	DD	HH	GF

TABLE III

SUCROSE LEVELS OF WATER AND NECTARS

(Experiment II)

Stimulus-set	% Sucrose increment: water	% Sucrose increment: nectar
I	5.0 vs. 5.5	5.0 vs. 5.5
II	5.0 vs. 5.25	5.0 vs. 5.3
III	5.0 vs. 5.0	5.0 vs. 5.0

TABLE IV

EFFECT OF COLOR ON RESPONSE

(Sucrose and apricot flavoring constant within each pair.)

		Os' opinion of colored solutions*							
		Untrained *Os*: Sweetness, 8 vs. 12% sucrose				Trained *Os*: Sweetness, 10 vs. 12% sucrose			
Illumination	Factor evaluated	No. judg.	% same	% more	% less	No. judg.	% same	% more	% less
Daylight	sweetness	101	**67.3**	21.8	10.9	80	**73.8**	13.8	12.4
	flavor	101	**62.4**	20.8	16.8	80	**70.0**	11.3	18.7
Red light	sweetness	87	**74.7**	12.7	12.6	80	**75.0**	11.3	13.7
	flavor	87	**65.4**	17.3	17.3	80	**52.5**	23.8	23.7

* Correct responses are set in boldface. All of these responses are significant at the 1/10% level.

TABLE V

EFFECT OF COLOR ON RESPONSE

(Greater apricot flavoring with sucrose-level constant within each pair.)

		O's opinion of solution with greater flavoring*							
		Untrained *Os*: Sweetness, 8 vs. 12% sucrose				Trained *Os*: Sweetness, 10 vs. 12% sucrose			
Illumination	Factor evaluated	No. judg.	% same	% more	% less	No. judg.	% same	% more	% less
Daylight	sweetness	104	**70.2‡**	18.3	11.5	44	**75.0‡**	11.4	13.6
	flavor	104	51.9‡	**31.7**	16.4	44	54.5†	**38.6**	6.9
Red light	sweetness	95	**84.2‡**	6.3	9.5	48	**81.3‡**	14.6	4.1
	flavor	95	60.0‡	**28.4**	11.6	48	41.7	**52.1†**	6.2

* Correct response is set in boldface. † Significant at the 1% level. ‡ Significant at the 1/10% level.

sucrose and flavoring were held constant within pairs; in Table V, sucrose was constant with greater flavoring in the first sample; and in Table VI, the flavoring was constant with more sucrose in the first sample.

When both sucrose and flavoring were constant within pairs, a significant number of correct responses was obtained regardless of illumination (Table IV). The untrained panel of *Os* was influenced slightly by the color-variation as seen by comparing the values for more sweetness (21.8%) and for more flavor (20.8%) ascribed to the colored solutions evaluated under daylight illumination. Under red illumination, the untrained panel's incorrect responses were distributed equally between more and less.

When the level of sucrose was constant between pairs, the *Os'* responses to sweetness showed a high degree of accuracy. There was, however, a significant number of incorrect judgments of flavor (Table V). Under red illumination, the untrained panel gave a more accurate report of sweetness. The *Os* of the trained panel were, apparently, distracted by color since their performance improved when solutions were tasted under red lights.

TABLE VI

EFFECT OF COLOR ON RESPONSE

(Greater sucrose level with apricot flavoring constant within each pair.)

Illumination	Factor evaluated	*Os'* opinion of solution with greater level of sucrose* — Untrained *Os*: Sweetness, 8 vs. 12% sucrose				Trained *Os*: Sweetness, 10 vs. 12% sucrose			
		No. judg.	% same	% more	% less	No. judg.	% same	% more	% less
Daylight	sweetness	104	5.8	**94.2**	0	40	5.0	**90.0**	5.0
	flavor	104	**47.1**	37.5	15.4	40	**77.5**	17.5	5.0
Red light	sweetness	94	8.5	**88.3**	3.2	48	8.3	**89.6**	2.1
	flavor	94	**58.5**	26.6	14.9	48	**43.8**	39.6	16.6

* Correct responses are set in boldface. All of these responses are significant at the 1/10% level.

Table VI shows that *Os* readily detected the variations in sweetness between paired samples regardless of illumination. Whereas untrained *Os* responded to flavor with greater accuracy under red illumination, the opposite was the case with trained *Os*.

Variation in individual performance explains, in part, the discrepancies observed in the averaged values of the trained panel. Under daylight illumination, accuracy ranged from 60 to 100% and from 31 to 75% for sweetness and flavor, respectively. Accuracy of the responses to sweetness and flavor when solutions were judged under red illumination ranged from 58

to 100% and from 31 to 75%, respectively. Although individuals were affected differently by illumination, the over-all totals indicate an increased accuracy for sweetness and a decreased accuracy for flavor as a result of the color-equalizing effect of red illumination.

Green coloring did not affect the accuracy with which judges identified sweetness and peppermint (Table VII). These samples were not evaluated under red illumination since this light did not mask the green coloring. The red lights did, however, cause a slight decrease in accuracy in the cherry-flavored solutions. Of the incorrect responses, a greater number ascribed more rather than less flavor and sweetness to the colored solutions.

The *Os* of the untrained panel were more accurate in their judgment of orange-colored, apricot-flavored solutions when color-differences were masked by red illumination (Table VIII). Since each sample was compared against every other sample an equal number of times, the response to the

TABLE VII

EFFECT OF COLOR ON JUDGMENTS OF SUCROSE AND PEPPERMINT OR CHERRY FLAVORING

(Sucrose and flavoring constant within each pair.)

		Os' opinion of colored solution*							
		Peppermint (Stimulus-set III)				Cherry (Stimulus-set II)			
Illumination	Factor evaluated	No. judg.	% same	% more	% less	No. judg.	% same	% more	% less
Daylight	sweetness	106	**84.9**†	8.5	6.6	100	**57.0**†	26.0	17.0
	flavor	106	**74.5**†	13.2	12.3	100	**59.0**†	22.0	19.0
Red light	sweetness	0	—	—	—	100	**51.0**	30.0	19.0
	flavor	0	—	—	—	100	**54.0**	28.0	18.0

* Correct responses are set in boldface. † Significant at the 1/10% level.

sweetness of each color was totalled and the values converted to percentages.

In aqueous solutions, there were no significant deviations from the expected 25% level, even when samples identical in sweetness were compared. This result seems to indicate that the panel did not associate sweetness with any specific color. In pear nectar, however, there was a pronounced tendency to designate the green-colored as the least sweet. This association was evidenced to a significant extent in the series of identical samples where decisions were made on the basis of color alone. It is possible that the panel associated tart tastes with green-colored foods and thereby ascribed less sweetness to green pear nectar. The greatest amount of sweetness was associated with the natural, uncolored nectar when samples of identical sugar content were compared. Totalling the 576 nectar com-

TABLE VIII

RESPONSE TO DUPLICATE SOLUTIONS IN PAIRED COMPARISON

(Response in terms of first sample within each pair with correct response are in boldface type. All responses significant at the 1/10% level.)

		Untrained *Os*				Trained *Os*											
		apricot				peppermint				cherry				apricot			
Illumination	Factor evaluated	No. judg.	% same	% more	% less	No. judg.	% same	% more	% less	No. judg.	% same	% more	% less	No. judg.	% same	% more	% less
Daylight	sweetness	37	**62.1**	20.3	17.6	49	**97.8**	1.1	1.1	31	**64.4**	19.3	16.3	24	**83.4**	8.3	8.3
	flavor	37	**59.4**	26.4	14.2	49	**73.4**	13.9	12.7	31	**64.4**	19.3	16.3	24	**50.0**	37.5	12.5
Red lights	sweetness	41	**83.0**	9.0	8.0	0	—	—	—	32	**59.3**	35.1	15.6	24	**75.0**	20.8	4.2
	flavor	41	**75.6**	12.2	12.2	0	—	—	—	32	**65.6**	15.6	18.8	24	**58.3**	29.2	12.5

parisons showed green nectar to be least sweet and red nectar to be sweetest.

According to Ryan, explaining results like those described above as due to association or conditioning does more harm than good as the implication is that the problem is solved.[9] We agree with Ryan that the study of coöperative functions of the senses should involve complete knowledge of all the dynamic and causal properties of the perceived objects. Further investigations and new techniques of study are needed for a better understanding of the role of color in taste discrimination, food selection, and ultimate consumption.

TABLE IX

PERCENTAGE OF SWEETNESS RESPONSES IN COLORED AND COLORLESS SOLUTIONS (Water and Pear Nectar)

Sucrose increment (%)		Color of solutions			
		colorless	red	green	yellow
	(expected)	25.0	25.0	25.0	25.0
WATER:	No. judg.				
5.0 vs. 5.5	144	23.6	26.4	22.2	27.8
5.0 vs. 5.25	288	27.8	20.5	25.3	26.4
5.0 vs. 5.0	144	20.8	22.9	25.0	31.3
Total	576	25.0	22.6	24.5	27.9
NECTAR:					
5.0 vs. 5.5	144	22.9	31.3	22.2	23.6
5.0 vs. 5.3	288	24.0	25.0	23.3	27.7
5.0 vs. 5.0	144	33.3*	25.7	16.0*	25.0
Total	576	26.0	26.7	21.2	26.1

* Significant at 5% level.

SUMMARY

The effect of food coloring on the sweetness and flavor of various aqueous and nectar solutions was determined by the method of paired comparison. Green coloring did not affect the accuracy with which *O*s identified sucrose level and peppermint flavoring in aqueous solutions. Untrained panels had a slight tendency to ascribe greater sweetness and greater flavor to orange- and red-colored solutions containing apricot and cherry flavoring, respectively, especially when sucrose differences were very small. When red illumination was used to mask color differences, the un-

[9] T. A. Ryan, Interrelations of the sensory systems in perception. *Psychol. Bull.*, 31, 1940, 691.

trained *Os* were more accurate whereas the highly trained *Os* were apparently distracted and showed a decreased accuracy. Both panels showed tendencies to ascribe more sweetness and more flavor to the first sample within a pair.

Sweetness-discrimination was not influenced by red, green, or yellow coloring in unflavored aqueous solutions. In pear nectar, there was a pronounced tendency to designate the green-colored samples as the least sweet. It is speculated that the green color was associated with tartness or lack of sweetness.

SYMBOLIC ACTIVITY IN 'LEARNING WITHOUT AWARENESS'

By SHERMAN J. TATZ, Adelphi College

A number of recent studies have been at least in part concerned with the issue of 'learning without awareness.'[1] The original experiment in this area was performed by Thorndike and Rock,[2] and was used by these authors to support Thorndike's hypothesis that reward acts directly and automatically to strengthen stimulus-response connections.

Thorndike and Rock's criterion of learning without symbolic mediation, however, was subsequently discredited by Irwin, Kaufman, Prior, and Weaver.[3] The issue was recently re-opened by a study by Greenspoon.[4] His *Ss* were asked to say single words as they came to mind in a free-association setting. Greenspoon had nine different groups in all, three of which are particularly relevant here.[5] In one group he reinforced all plural nouns by saying "mmm-hmm" after each such word; in another group he reinforced in the same way all words except plural nouns; a control group was given the same task but no reinforcement was used. Compared with the control group, the frequency of the reinforced response increased in both of the former groups even though subsequent questioning of the *Ss* failed to reveal awareness of the relation between their responses and the reinforcement. While Greenspoon did not relate his results to the hypothesis of direct action of rewards, others have drawn the implication that his results demonstrate the direct and unconscious effect of reward.[6]

Subsequent investigators have reported similar findings. Sidowski used a procedure essentially the same as Greenspoon's, but employing only a light as reinforcement.[7] Cohen, Kalish, Thurston, and Cohen used a somewhat different pro-

---

* Received for publication March 14, 1959. This article is based on a dissertation submitted in partial fulfillment of the Ph.D. degree to the Department of Psychology of Yale University. The study was directed by Dr. Fred D. Sheffield.

[1] Such studies are reviewed by J. K. Adams, Laboratory studies of behavior without awareness, *Psychol. Bull.*, 54, 1957, 383-405; and by Leonard Krasner, Studies of the conditioning of verbal behavior, *Psychol. Bull.*, 55, 1958, 148-170.

[2] E. L. Thorndike and R. T. Rock, Jr., Learning without awareness of what is being learned or intent to learn it, *J. exp. Psychol.*, 17, 1934, 1-19.

[3] F. W. Irwin, K. Kaufman, G. Prior, and H. B. Weaver, On 'learning without awareness of what is being learned,' *J. exp. Psychol.*, 17, 1934, 823-827.

[4] Joel Greenspoon, The effect of verbal and non-verbal stimuli on the frequency of members of two verbal response-classes, unpublished Doctoral dissertation, Indiana University, 1951; The reinforcing effect of two spoken sounds on the frequency of two responses, this JOURNAL, 68, 1955, 409-416.

[5] Greenspoon, Dissertation.

[6] John Dollard and N. E. Miller, *Personality and Psychotherapy*, 1950, 43-44; J. A. McGeoch and A. L. Irion, *The Psychology of Human Learning*, 2nd. ed., 1952, 269-271.

[7] J. B. Sidowski, Influence of awareness of reinforcement on verbal conditioning, *J. exp. Psychol.*, 48, 1954, 355-360.

cedure.[8] *S* was asked to make up a sentence using each of a series of verbs and beginning with any one of six personal pronouns. As reinforcement *E* said "good" after each sentence beginning either with 'I' or 'we.' In both of the above studies, learning was reported with *Ss* who, on subsequent questioning, did not correctly state the principle according to which reinforcements were allotted.

The significance of such studies for the Thorndikian hypothesis of direct action of reward depends on the assumption that 'awareness' was absent, the implication being that correct symbolic representation by *S* of the conditions of reinforcement would provide an alternative, indirect, mechanism for the action of reward in producing an increase in frequency of correct responses. The authors of the more recent studies cited above have excluded the data of *Ss* who were able to verbalize *correctly* the system of reinforcement,[9] but they apparently have not taken account of *partially* correct systems, *i.e.* systems which could mediate a level of performance higher than chance even though they did not correspond exactly to *E*'s system of allotting reinforcements. The purpose of the present study was to investigate, in an experimental situation comparable to those in the studies cited above, the relation between *S*'s verbalizations—whether correct or not—and the acquisition of 'correct' responding.

An operant responding procedure was employed, using numerical responses and selecting particular digits to be reinforced. This provided discrete responses that presumably could be acted upon more directly by the reward, as contrasted with the complex classes of responses reinforced by Thorndike and Rock and by Greenspoon.[10] 'Good' was used as the social-approval reward rather than Greenspoon's ambiguous 'mmm-hmm,' and as a control 'reinforcement' a clearly audible click was used instead of 'good' with some *Ss*. Instructions calculated to have differential effects in motivating *S*'s attempts to discover the basis for 'reinforcement' were employed, and an intensive interview was used at the end to determine the content of the *S*'s implicit verbalizations during the experimental procedure.

METHOD

Experimental design. Four groups of *Ss* were studied, with 20 *Ss* in each group. The *Ss* were studied in 10 successive replications of 8, with 2 *Ss*' being randomly assigned to each of the 4 groups in each replication. Five pairs of digits to be reinforced were selected: 0 and 5, 1 and 6, 2 and 7, 3 and 8, and 4 and 9. Each of

[8] B. D. Cohen, H. I. Kalish, J. R. Thurston, and Edwin Cohen, Experimental manipulation of verbal behavior. *J. exp. Psychol.*, 47, 1954, 106-110.

[9] Cohen, Kalish, Thurston, and Cohen, *op. cit.*, 106-110; Sidowski, *op. cit.*, 355-360; Greenspoon, Dissertation.

[10] Thorndike and Rock, *op. cit.*, 1-19; Greenspoon, Dissertation.

these pairs was randomly assigned to two of the replications as the 'correct' digits within a given replication. Each *S* said a three-digit number every time a signal light came on, varying the digits each time. Any number containing either or both of the 'correct' digits was followed by the 'reinforcement' appropriate to that group. The experimental groups used were as follows:

'Good'—no information group (G-NI). The reinforcement consisted of *E*'s saying 'good' with a rising inflection, as if to indicate that *S* was following instructions or that *E* was keeping up with *S*. No information was given about the existence of 'correct' responses.

Click—no information group (C-NI). The reinforcement consisted of an audible click produced by *E*. No information was given about the existence of 'correct' responses.

'Good'—informed group (G-I). The reinforcement was the same as for Group G-NI. At the beginning, however, additional instructions were given to the effect that some numbers were correct, that correct ones would be followed by 'good,' and that *S* was to try to get a 'good' every time.

Click—informed group (C-I). The reinforcement was the same as for Group C-NI. At the beginning, however, additional instructions were given as in Group G-I.

In all groups, following the experimental procedure, *S* was interviewed to get at his conception of the experiment's purpose, conditions, and so forth.

Subjects. The *Ss* were 80 women in the introductory course in psychology at Connecticut College, New London, Connecticut.

Apparatus. S sat at a table with her back to *E*, who sat at a different table. Immediately before *S* was a tall, white screen, to one side of which (visible to *S*) was a signal-light: a 15-w. frosted lamp. Hidden from view was a sound-deadened box which contained the motor controlling the signal-light. When the signal-light was in use it went on for 1 sec. every 6.66 sec. Attached on the underside of *E*'s table and hidden from *S* were an electro-magnetic counter and a switch which operated it. Momentarily closing the switch produced a resounding click.

Procedure. E gave the following basic instructions to all the *Ss* (with exceptions noted below).

At this point I'll explain only what I want you to do. But when you're all finished we can discuss the experiment, and I'll tell you what it's all about. Every 7 sec. this light will come on. [*E* turned on the switch activating the signal light.] Each time it comes on, I want you to say a three-digit number, that is, any combination of three numbers, each one being anything from zero to nine. Say them in mixed-up order, just as they come into your head when the light comes on. Do you want to try a few with the light now, just to make sure you've got the idea? [*S* was allowed to make two or three responses, *E* insuring that *S* knew that each response was to be said as three separate digits, *e.g.* 'six, five, nine' rather than 'six hundred and fifty nine.'] Remember, you can use zero, too. Now I'll tell you when to begin, and you keep on going until I tell you to stop.

With Groups G-I and C-I, instead of saying, "Say them in mixed-up order . . ." *E* said, "Start out saying them in mixed-up order . . ." With these groups, too, the following additional instructions were substituted for the last sentence above:

Now, as I said, start out saying the numbers just as they come into your head when the light comes on. But according to a pre-arranged key that I have, cer-

tain three-digit numbers are correct and others are not. [At this point Group G-I was told] "Whenever you say one that's correct, I'll say 'good.' I wish you to learn to say numbers so correctly that I will say 'good' every time. I'll tell you when to begin, and you keep on going until I tell you to stop." [At the corresponding point, Group C-I was told] Whenever you say one that's correct, you'll hear this click [demonstrating]. I want you to learn to say the numbers so correctly that you hear the click every time. I'll tell you when to begin, and you keep on going until I tell you to stop.

E then took his seat and told *S* to begin. The experimental procedure continued with *E* presenting the prescribed reinforcer at the appropriate times, until *S* had made 160 three-digit responses. All responses were recorded on a data sheet. *S* then was interviewed.

Content of interview. There were two forms of the questionary upon which the interviews were based: one for Groups G-NI and C-NI, and the other for Groups G-I and C-I. The first nine questions were the same in both forms, but the remaining questions differed somewhat to suit the different experimental groups. The questions proceeded from general to specific, giving *S* an initial chance to volunteer information indicating 'awareness' but later asking directly about this matter. The procedure involved much lengthier probing than had been used to detect 'awareness' in previous experiments, and *S* had a better opportunity to recall and put into words any relevant hypotheses she may have had during the experiment.[11]

Results

Over-all effects. The differential effects of the experimentally manipulated variables on performance of the critical overt behavior are shown in Fig. 1. A correct response is defined as a three-digit number including either or both of the digits selected to be reinforced for any *S*. It can be seen in the figure (*a*) that Groups G-I and C-I were essentially the same in rate and level of acquisition; (b) that Group G-NI showed definite acquisition, but at a level lower than Groups G-I and C-I; and (*c*) that Group C-NI appeared to show a decrease in correct responding.

The total correct responses of the four groups were subjected to an analysis of variance to test the significance of the over-all effects. It was found that the informed groups were superior to the uninformed groups ($p < 0.01$) and that 'good' over-all was superior to the click ($p < 0.01$) as a reinforcing stimulus, using the within-groups variance as the error-term in each case. Of still greater interest is the fact that the interaction between information and reinforcing stimulus was significant ($p < 0.02$). As evident in Fig. 1, while 'good' and click were equally effective reinforcements for the informed groups, 'good' produced acquisition whereas

[11] Both forms of the questionary appear in S. J. Tatz, Symbolic mediation in 'learning without awareness,' unpublished Doctoral dissertation, Yale University, 1956, Appendix, 1-3.

the click did not for the uninformed groups, an avoidance tendency even being present in Group C-NI.

Subsequent t-tests involving the above variables revealed the following: there was no significant difference between Groups G-I and C-I ($t = 0.07$, $p > 0.9$);[12] Group G-NI was superior to Group C-NI ($t = 3.92$, $p < 0.01$); Group G-I, to Group G-NI ($t = 2.81$, $p < 0.01$); and Group C-I to Group C-NI ($t = 7.76$, $p < 0.01$).

To test for the presence of acquisition, another set of t-tests was made

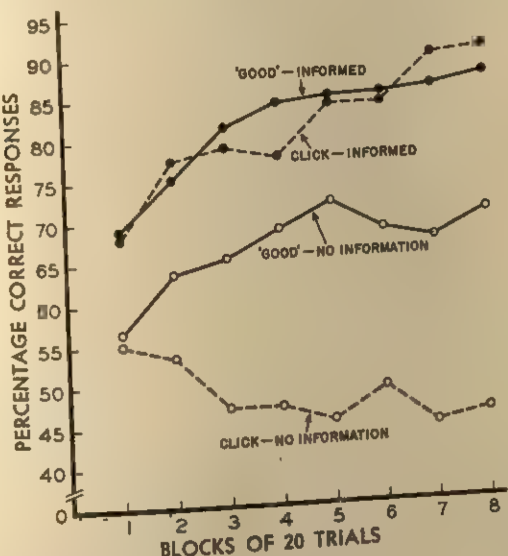


FIG. 1. PERFORMANCE OF THE FOUR GROUPS IN TERMS OF PERCENTAGE OF CORRECT RESPONSE IN EACH BLOCK OF 20 TRIALS.

of the mean difference in correct responses between the first and last blocks of 40 trials in all four groups. Acquisition was revealed in Group G-I ($t = 4.23$, $p < 0.01$), Group C-I ($t = 4.16$, $p < 0.01$), and Group G-NI ($t = 2.57$, $p < 0.02$); a decline was shown in Group C-NI ($t = 5.00$, $p < 0.01$).

Effects of experimental variables. In the interviewing which followed the experimental procedure, a considerable variety of kinds of 'awareness' was revealed. The *Ss* were classified according to the degree to which their

[12] In this and all subsequent t-tests, the p-value given is for both tails.

level of correct responding would be raised above chance if they performed in line with their verbalizations. From the varied contents of verbalization, four classes could fairly readily be discriminated:

Class 1. *S* verbalized correctly the basis on which reinforcements occurred, *i.e.* she named both of the correct digits, making it clear that a response containing either one in any of the three positions would be followed by a reinforcement.

Class 2. *S* described a system different from that used by *E* but which, if *Ss* were to follow it completely, would produce 100% reinforcement. *Ss* in this class generally learned one correct digit, but they sometimes simply memorized several successful three-digit combinations. Others had more elaborate partial solutions: for example, one *S* who was reinforced for 0 and 5 described the successful combinations as "ones divisible by 5."

Class 3. *S* described a system which, if adopted, would result in a level of correct responding above chance but below 100%. Generally *S* had acquired some relevant as well as irrelevant information or devised a complex pattern which increased the probability of reinforcement. For example, one *S* reinforced for 2 and 7 reported the digits which 'worked' as follows: "1, 7, 9, 5 or maybe combinations of them. Didn't always work individually."

Class 4. *S* verbalized no principle which would enable her to respond correctly above chance. Some of the *Ss* in this category had not even considered systems. Others reported attempts which would not have improved on chance, *e.g.* "Thought you said it whenever I said successive numbers."

TABLE I

NUMBER OF *Ss* IN EACH CATEGORY OF 'AWARENESS' BY EXPERIMENTAL GROUP

Category	Group			
	G-I	C-I	G-NI	C-NI
(1) Stated the system	2	3	0	6
(2) Stated a 'successful' system	13	12	8	2
(3) Stated a partially 'successful' system	5	5	2	4
(4) Stated no effective system	0	0	10	8

Table I shows the distribution of *Ss* among the four classes as a function of experimental condition. It is evident that, except for Group C-NI, very few *Ss* were able to describe the actual basis for reinforcing. It will be noted also that all of the informed *Ss* arrived at a system that was at least partially successful, and that only 18 of the 80 *Ss* remained completely naïve with regard to the basis for reinforcement.

To ascertain the motivation with respect to obtaining reinforcement of *Ss* in the uninformed groups, a tabulation from information in the interview protocols was made which characterized *Ss* according to the following three motivational classes:

Class A. *S* stated she tried to get the reinforcer on every trial.

Class B. *S* stated she tried to get the reinforcer only part of the time, as when testing hypotheses as to correct responses, or that she did not try to get or avoid the reinforcer. For example, one *S*, when asked whether she tried to make the 'reinforcer' come every time, said, "No, nor avoid. Tried both at different times, just to see what the reaction would be."

Class C. *S* stated that she tried to avoid bringing the reinforcer or that it was annoying. For example, one *S* said, "The click was sort of annoying—it came with those numbers. I thought I shouldn't use them. Bad numbers."

TABLE II

NUMBER OF *Ss* IN EACH MOTIVATIONAL CATEGORY IN THE UNINFORMED GROUPS

Category	Group	
	G-NI	C-NI
(A) Attempted to get 'reinforcer' every time	8	1
(B) Attempted to get 'reinforcer' part of the time	10	10
(C) Attempted to avoid 'reinforcer,' or found 'reinforcer' annoying	2	9

$X^2=9.9$, $df.=2$, $p<0.01$.

TABLE III

PERCENTAGE OF CORRECT RESPONSES OVER ALL 160 TRIALS IN EACH CATEGORY OF AWARENESS FOR EACH EXPERIMENTAL GROUP

Category	Group			
	G-I	C-I	F-NI	C-NI
(1) Stated the system	59.7	77.3	—	40.2
(2) Stated a 'successful' system	90.1	90.5	84.6	45.7
(3) Stated a partially 'successful' system	68.5	61.5	60.0	55.2
(4) Stated no effective system	—	—	53.5	51.7

Table II shows the distribution of *Ss* among these classes for the two uninformed groups. It can be seen that these *Ss* in general verbalized a tendency to attempt to bring 'good' every time, but to avoid the click.

Relations between correct responding and symbolic behavior. Table III shows the mean percentage of correct responses, over all 160 trials, of the experimental groups broken down by the symbolic categories used in Table I. The over-all *F* for the 13 cells in which cases appear is 20.96 (significant well beyond the 1% level). The following features of the results shown in Table III should be mentioned.

(*a*) In both Groups G-I and C-I, the few *Ss* who 'caught on' to the complete system (Class 1) responded at a lower level than those (Class 2) who had a successful system but did not learn the entire system used by *E*. The difference between Class 1 and Class 2 was significant for Group G-I ($t = 4.62$, $p < 0.01$) but not for Group C-I ($t = 1.74$, $p > 0.10$).

(*b*) In Groups G-I, C-I, and G-NI, those who had partially successful systems

(Class 3) were inferior to those who had 'incorrect' but completely successful systems (Class 2). The t for combined *Ss* of Class-3 versus combined those of Class-2 in these groups was 6.12 ($p < 0.01$).

(*c*) Those in Groups G-I, C-I, and G-NI who had partially successful systems (Class 3) were superior to those in Group G-NI having no effective system (Class 4), the t for combined *Ss* of Class-3 versus those of Class-4 in Group G-NI being 2.43 ($p < 0.05$).

(*d*) The *Ss* of Group C-NI in the four categories responded at levels averaging at or below chance, the lowest level being that of *Ss* who verbalized the complete system (Class 1). The t for the six *Ss* of Class 1 in this group versus the *Ss* of Class 4 was 3.23 ($p < 0.01$).[13]

Discussion

The results demonstrate that level of correct responding is related to implicit symbolic behavior as determined by interview. The *Ss* with no effective system performed essentially at chance. The exact theoretical chance-level of correct responding cannot be calculated because of various non-random habits involved in guessing numbers.[14] If the *Ss* had actually started guessing at random, the probability of receiving a reinforcement on any given trial would be $1.0 - (8/10)^3 = 0.488$, or 48.8%. To the extent that *Ss* used the common non-random procedure of never repeating a digit within a given combination of three, this probability is raised toward $1.0 - (8/10)\ (7/9)\ (6/8) = 0.533$ or 53.3%. In any case, the performance of the *Ss* of Class 4 did not show any significant variation from beginning to end of the 160 responses.

In general, for Groups G-I, C-I, and G-NI, the more potentially effective the system verbalized, the higher the level of responding, whereas generally in Group C-NI, potentially more effective systems were related to lower levels of responding. This inverse relationship in Group C-NI is explained by the tendency, reported by *Ss* in this group, to avoid the click (see Table II). With motivation to *get* the reinforcement, the more effective the system used, the more frequently reinforcement was obtained; with motivation to *avoid* the reinforcement, the more effective the system used, the more effectively the reinforcement was avoided.

An exception to the relationships described is the case of the five informed *Ss* who verbalized *the* correct system but nevertheless performed at a level lower than that of *Ss* of Class 2. Their lower performance levels

[13] The pooling of data from *Ss* in different experimental conditions within the same category of 'awareness' was justified in part by the fact that an *F*-test for the cells involved in a given combination was in no case at all significant.

[14] W. O. Jenkins and Leta Cunningham, The guessing-sequence hypothesis, the 'spread of effect,' and number-guessing habits, *J. exp. Psychol.*, 39, 1949, 158-168.

can be explained in part by the fact that *Ss* who discover the complete system are likely to be those who have spent more time experimenting with various response-systems and therefore obtaining fewer reinforcements than *Ss* who discover and immediately utilize an incorrect system which nevertheless brings a high frequency of reinforcement. This explanation was borne out in the interview protocols of these *Ss*. One *S*, for example, said that after discovering two correct digits, she looked for a third. Another stated that after discovering the correct digits she looked for a pattern of digits that might also bring the reinforcement.

The implication of these findings is to cast doubt on the evidence for the direct action of reward in all experiments on 'learning without awareness' in which awareness has been treated as an all-or-none phenomenon. Although the importance of partial solutions was hinted at in an observation of Postman and Jarrett,[15] it does not appear to have been taken fully into account in a subsequent study by Philbrick and Postman,[16] employing a similar procedure. The latter authors concluded that awareness follows after a certain amount of acquisition on the basis of the automatically-acting reward, but they treated awareness as an all-or-none affair and based their conclusion on *Ss* who showed slight improvement before their ability to verbalize correctly. To the extent that the present experiment is representative of the previous studies in this field, the findings lead us to question the adequacy of the dichotomous criterion of awareness used in those studies and demonstrate that symbolic representation of the conditions of reinforcement is related to overt performance.

SUMMARY

The possible role of implicit symbolic activity in so-called 'learning without awareness' was studied in an operant situation. The nature of the 'reinforcement' and the instructions to *S* were varied, and the extent of *S*'s 'awareness' was probed in an intensive interview. The results were interpreted as indicating that previous investigators had probably not adequately ascertained possible effects of partially correct solutions, which may mediate performance changes in such situations.

[15] Leo Postman and R. F. Jarrett, An experimental analysis of 'learning without awareness,' this JOURNAL, 65, 1952, 244-255.

[16] E. B. Philbrick and Leo Postman, A further analysis of 'learning without awareness,' this JOURNAL, 68, 1955, 417-424.

THE INTERACTION OF SPACE AND TIME IN CUTANEOUS PERCEPTION

By Betty A. Wieland, University of California, Los Angeles

There have been many indications in the literature that temporal patterning of cutaneous stimuli has an influence on spatial perception. In 1855 Czermak reported that the simultaneous two-point threshold was larger than the successive two-point threshold.[1] Apparent movement, or the *Phi* phenomenon, and the *Tau* phenomenon reported by Helson and King are examples of the types of spatial temporal interaction that have been investigated.[2]

Recently Jones suggested that the error of localization and the simultaneous two-point threshold were two extremes of a continuum representing the interaction of space and time.[3] He obtained the two-point threshold on 2 *S*s using five temporal intervals between stimulation ranging from 0.1 to 1000 m.sec. and found that as the temporal interval increased the size of the two-point threshold decreased. Scherrer, Barbizet, and Hénaff also found this relationship as an incidental result of their studies of the threshold of discontinuity.[4]

Although no modern neurological theory predicts the space-time interactions which have been found in psychophysical studies, recent evidence does admit such a possibility. There does not appear to be a strict point-for-point representation of the periphery in either the somatic areas of the cortex or in the thalamus. Rather, a given peripheral point evokes activity over a considerable central area, and a given central locus responds to stimulation from a sizeable peripheral area. There are differences in evoked potentials from different points of peripheral stimulation, how-

* Received for publication May 4, 1959. This paper is based on a Doctoral dissertation directed by Professor F. Nowell Jones and submitted to the University of California at Los Angeles. The author is now at the University of Houston.

[1] E. G. Boring (*Sensation and Perception in the History of Experimental Psychology*, 1942, 477) reviews the early work on the successive two-point threshold.

[2] W. S. Hulin, An experimental study of apparent tactual movement, *J. exp. Psychol.*, 10, 1927, 293-320; Harry Helson and S. M. King, The *Tau* effect: An example of psychological relativity, *ibid.*, 14, 1931, 202-217.

[3] F. N. Jones, Space-time relationships in somesthetic localization, *Science*, 124, 1956, 484.

[4] Jean Scherrer, Jacques Barbizet, and M. O. Hénaff, Contribution à l'étude de l'organization temporo-spatiale de la sensibilité somesthétique normale, *Année psychol.*, 54, 1954, 53-81.

ever, not only in size but also in latency.[5] Amassian, recording from the somatic association area, found that single units responded to stimulation from a wide peripheral area but that the responses from different areas differed in shape.[6] The form of the response to a given stimulus could be altered considerably by the presentation of another stimulus preceding it by a short period of time. Complete blocking of the second stimulus could occur with interstimulus intervals up to 500 m.sec. or more. His results indicate the possibility of conversion of spatial intervals into temporal intervals within the central nervous system.

If time does in some way replace space in our perceptual mechanisms, it should be possible to vary perceived space in a spatial discrimination merely by varying the temporal interval in the stimulus-pattern. Further, it should be possible to determine the function relating the introduced time to the displaced space. This study was an attempt to investigate such a relationship using a variation of the successive two-point threshold. The problem was to have an observer judge whether one or two places had been stimulated when, with the distance between two stimuli held constant, the time between the stimuli was varied.

Method and Procedure

Observers. Data from five *Os* are presented. Three of the *Os* (*BE, RS,* and *JW*) were men and two (*JB* and *ST*) were women. BE, RS, and JB were graduate students in psychology; *ST*, an undergraduate student in psychology; and *JW*, the only non-student among the *Os*, had had undergraduate courses in psychology.

Equipment. Two model S4B Grass square-wave stimulators were used to deliver the stimuli. They were so wired that activating the first stimulator delivered a single pulse of 0.5-m.sec. duration through one electrode and at the same time so activated the second stimulator that, after a pre-set delay, the second stimulator delivered a 0.5-m.sec. pulse through another electrode. The delay could be varied from 0.01 to 1000 m.sec. The outputs of the two stimulators were led through a switching device which allowed either stimulator to fire either of the two electrodes independently. The active electrodes used were wick-type, silver-silver chloride electrodes. The area of the wick in contact with the skin was approximately 1 sq. mm. The inactive electrode was a 15 × 20-cm. copper screen covered with cheese-cloth and saturated with NaCl solution. The ventral surface of *O*'s right forearm rested on the inactive electrode. The active electrodes were held above the dorsal surface of the forearm

---

[5] W. R. Adey, I. D. Carter, and R. Porter, Temporal dispersion in cortical response, *J. Neurophysiol.*, 17, 1954, 165-181; S. M. Cohen and Harry Grunfest, Thalamic loci of electrical activity initiated by afferent impulses in cat, *idem.*, 193-207; R. M. Gaze and G. Gordon, The representation of cutaneous sense in the thalamus of the cat and monkey, *Quart, J. exp. Physiol.*, 39, 1954, 279-304.

[6] V. E. Amassian, Studies on ogranization of a somesthetic association area, including a single unit analysis, *J. Neurophysiol.*, 17, 1954, 39-47.

by a rack system so constructed that the position of the electrodes could be varied in three planes to obtain the desired positioning. One electrode was placed approximately 4 cm. above the styloid process and centered on the dorsal surface. The position of the second electrode was varied to obtain the different distances used, but was always approximately on the center line and proximal to the first electrode.

The input voltage to each electrode was adjusted until *O* reported that a sensation of pressure, well above threshold, was elicited. This voltage could be varied in two ways: by the voltage-output control of the stimulator, and by a resistance-control in the switching mechanism.

Pre-training. Since a large amount of time was required of each *O*, the *O*s were chosen on the basis of their availability and willingness to serve. This resulted in considerable variability in *O*s' knowledge about and experience with the phenomenon under investigation. The purpose of the experiment had been obvious to two pilot *O*s who were given only minimal information necessary for making the required judgments. In view of this finding and in an attempt to bring all the *O*s to a comparable level of knowledge concerning the experiment, all the *O*s were told the purpose of the experiment. Some discussion of cutaneous space-perception was held with each of them.

All the *O*s were given at least one session of about an hour of practice in the types of discrimination to be made. Those *O*s who had had no previous experience in this type of sensory discrimination were given more than one session. During these training sessions, the *O*s were given practice in distinguishing pressure from pain; in judging two stimuli for intensity and quality and in the determination of the two-point threshold. Results from pilot *O*s indicated that some difficulty was experienced in distinguishing between differences in time and differences in place. All the *O*s were, therefore, given some practice in judging differences in time between two stimuli as distinct from differences in space. It was emphasized during these trials that this was *not* the type of judgment that they were to make during the actual experiment. Training in these judgments was carried out in the same room and under the same conditions as the experiment itself. Training was continued until the *O*s reported that they were confident of their judgments and until, in *E*'s estimation, they were giving consistent judgments. They were then given a pre-trial with the experimental procedure. This was an abbreviated session in that only two or three distances between electrodes were used and only an ascending or a descending threshold was obtained for each distance. When *O* reported that he understood the type of judgment that he was to make and that he was reasonably confident of his judgments, the training period was terminated.

Procedure. O was seated in a chair beside a table upon which the equipment was placed. *E* sat on the other side of the table. *O* placed his right arm through a hole in a screen which hid his arm, the equipment, and *E* from his view. The arm was so placed that the volar side of the forearm rested on the inactive electrode. His position was adjusted until *O* judged it to be one in which he could sit comfortably for some time without moving. *O* was then given the following instructions.

Instructions. You are to judge whether one or two locations on your skin have been stimulated. Sometimes you will feel two stimuli definitely separated in time, sometimes you will not. You are to ignore the time-separation and judge only whether two places have been stimulated.

The active electrodes were positioned at a pre-selected distance apart. Seven distances (center to center of the electrodes) were used in each series: 1.0, 1.6, 2.4, 3.7, 5.9, 9.3, and 14.6 cm. The order of the distances in each series was selected by a random process. After the electrodes had been positioned, the proximal stimulus was presented and adjusted until *O* stated that it met the criteria which he had established as to quality and intensity, then the distal stimulus was presented and similarly adjusted. Finally both stimuli were presented, separated in time by 1 sec., and any adjustments necessary to make them equal in quality and intensity were made.

An attempt was always made to obtain pure sensations of pressure without an element of prick or pain. This frequently required the relocation of one or both electrodes. They were always kept the same distance apart, however, and on an approximate longitudinal axis of the arm. Frequent readjustments were necessary to keep the stimuli judged equal in quality and intensity.

After the two stimuli were judged equal in quality and intensity, the two-point threshold was obtained by varying the time between the two stimuli, using the method of limits. The proximal position was always stimulated first, followed by the distal position. The threshold was defined as that temporal separation at which *O* changed his judgment from 'one place' to 'two places' (in ascending series) or from 'two places' to 'one place' (in descending series). The size of the temporal steps used varied with the range of temporal intervals, ranging from 0.05 m.sec. at intervals less than 0.5 m.sec. to 50 m.sec. at intervals greater than 500 m.sec. Occasionally, as a control, only one place was stimulated.

After each threshold was obtained, *O* was given a period of rest, usually 2 or 3 min., but once or twice during the session, 5 to 10 min. For *BE*, *RS*, and *JB*, one ascending and one descending threshold at each distance were obtained in each session. For *JW* and *MS*, two ascending and two descending thresholds were obtained. A total of 12 thresholds was obtained for each distance, requiring 6 and 3 sessions, respectively.

RESULTS

The data represent the time-intervals at which *O* changed his judgment from 'one place' to 'two places' or from 'two places' to 'one place.' They were first treated by an analysis of variance to determine whether there were differences in this time-interval at different distances (*i.e.* spatial separation of stimuli) and with different *O*s. Since it was obvious from the raw data that the variances were not homogeneous, a log transformation was first employed and the analysis of variance was applied to the transformed data. For all but one of the *O*s the distance of 14.6 cm. was greater than the simultaneous two-point threshold. The data for this distance were not included in the analysis. Table I gives the summary of this analysis. One trial for one *O* at one distance was missing and was estimated. The total degrees of freedom and the within-cells degrees of freedom were reduced by one as a result of the missing datum. The

within-cells variance was used as the error-term to test the interaction-effect. Since the interaction was found to be significant beyond the 1/10% level of confidence, the interaction-variance was used as the error-term in testing the main effects. Both main effects were significant. The hypothesis that there is no difference in the time-interval at which *O* changed his judgment from 'one' to 'two' or from 'two' to 'one' at different spatial separations is rejected at the 1/10% level of confidence; the hypothesis that there is no difference in this time-interval with different *O*s is rejected at the 5% level of confidence.

TABLE I

SUMMARY OF THE ANALYSIS OF VARIANCE OF THRESHOLD SCORES EXPRESSED AS THE LOGARITHM OF TIME

Source	Sum of squares	*df.*	Var. est.	*F*
*O*s	27.7867	4	6.9467	3.1751*
Distances	279.1172	5	55.8234	25.5146†
D×*O*	43.7587	20	2.1879	15.5612†
Within cells	46.2434	329‡	0.1406	
Total	396.9060	358‡		

* $P<0.05$. † $P<0.001$. ‡ See text for explanation.

TABLE II

EQUATIONS FOR CURVES FITTED TO MEAN TIMES AT EACH DISTANCE FOR EVERY *O* AND THE INDICES OF DETERMINATION

O	Equation*	Index of determination $=p^2$†
BE	$\log T=2.83-.129\,D$	.937
JW	$\log T=3.29-.307\,D$	.941
RS	$\log T=2.94-.288\,D$	.959
JB	$\log T=2.94-.203\,D$	.963
ST	$\log T=3.28-.413\,D$	.963

* T = time between stimuli; D = distance between stimuli.
† p^2 = the square of the product moment correlation coefficient between $f(T)$ and $f(D)$.

The significant interaction indicates that the relationship between space and time is different for different *O*s. It was concluded, therefore, that the data from each *O* should be treated separately. Curves were fitted to the data for each *O* relating the spatial interval and the mean temporal interval of the two-point threshold. The equations for the best fitting curves using a least squares criterion are given in Table II as well as the index of determination (p^2), *i.e.* the square of the product moment correlation coefficient between $f(T)$ and $f(D)$. Table III gives the mean time-values obtained at each distance for each *O* and the values for these distances predicted from the equations.

Discussion. The results of the present study are in agreement with those of previous studies showing that time and distance can interact in the perception of cutaneous space. In addition, they show the relationship which space and time may have under one limited set of conditions. At very small spatial separations a relatively long time-interval was necessary for two stimuli to be judged as occurring in two places, while at greater spatial separations a shorter time-interval sufficed. The equations derived to describe the curves which appear to fit the obtained data imply that distance is an exponential function of time or, conversely, time is a logarithmic function of distance. A simple transformation of the general equation obtained permits another way of looking at the results. The equation $\text{Log } T = a - bD$ may be written $(T)\ (10^{bD}) = 10^{a}$. This states

TABLE III

PREDICTED AND OBTAINED VALUES OF LOG TIME

		Distance					
O		1.0	1.6	2.4	3.7	5.9	9.3
BE	log T'*	2.70	2.62	2.52	2.36	2.07	1.63
	log T†	2.73	2.58	2.57	2.42	1.87	1.72
JW	log T'	2.99	2.80	2.56	2.16	1.48	.44
	log T	2.87	2.86	2.53	2.01	1.93	.22
JB	log T'	2.73	2.61	2.45	2.19	1.74	1.05
	log T	2.80	2.52	2.32	2.37	1.74	1.02
RS	log T'	2.66	2.48	2.25	1.88	1.24	.26
	log T	2.83	2.26	2.29	1.73	1.50	.17
ST	log T'	2.87	2.62	2.29	1.75	.85	−.56
	log T	2.80	2.66	2.06	1.80	1.29	−.80

* Log T' = predicted value of log time; † Log T = obtained value of log time.

that the product of time and a function of distance is a constant. In other words, when the product of time and a certain function of distance reaches a certain constant value, for a given individual, the individual begins to discriminate between 'one' and 'two.' Within certain limits, space and time appear to be interchangeable.

The phenomenon of spatial temporal interaction in this type of two-point threshold is, without doubt, related to the *Tau* phenomenon of Helson and King.[7] They also found that judgments of spatial extent could be altered by altering the temporal relations between stimuli. The methodology used in the present study expresses in spatial-temporal terms the absolute threshold of spatial discrimination, while Helson and King's methodology yields the *JND*. It is possible that the parameters in the

[7] Helson and King, *op.cit.*, 202-217.

equations expressing how time may replace space would be the same with the two methods, but it seems more likely that they would differ with the size of the standard stimulus (*i.e.* the standard spatial separation). One might expect to find a Weber fraction here. The published data of Helson and King do not lend themselves to such an analysis and, since they used mechanical stimulation, their results cannot be compared to the present ones using electrical stimulation.

Although the general forms of the curves relating space and time appear to be similar for different *O*s, the parameters describing these curves differ. The differences might be attributed in part to basic structural differences between *O*s such as differences in length of limbs and length of pathways. Part of the differences are probably due to differences in criteria of judgment adopted by the several *O*s. It is certainly not unusual to find individual differences between *O*s in sensory experiments. Where relatively complex judgments such as those required here are involved, it is to be expected. When further research gives evidence for the factors which change the values of the parameters within one *O*, we may be able to understand better the source of individual differences.

It is tempting to speculate on the neural mechanisms involved in this spatial temporal interaction. Jones has suggested that we should look to the association areas investigated by Amassian.[8] The correspondence is suggestive. Blocking of a second impulse by a first impulse from two stimulations of the same area (in the cat) occurs approximately through the time-span defining the error of localization in man. Two stimuli applied to different points activate the same area but with different temporal patterns which might well be altered by changes in temporal patterning of the stimuli. We might also look to the somatic areas I and II and even the thalamus where we have indications of overlapping of areas excited by different peripheral points but where maximal potential size as well as latency appears to differentiate between peripheral sites of stimulation. The neural processes, however, underlying the interaction of space and time must certainly be complex and cannot be expected to be related in any simple manner to the results of recording isolated cortical responses.

Although we cannot specify the exact neurophysiological mechanisms underlying this interaction-phenomenon, the psychophysical results are not at odds with the neurophysiological evidence concerning the mechanisms of space-perception. Both methodologies point to a spatial temporal

[8] Jones, *op.cit.*, 484; Amassian, *op.cit.*, 39-47.

gradient as underlying space-perception. The fact that the same general form of equation appears to fit the data from different individuals gives evidence for the reliability of the form of the function under the conditions used in this experiment. It remains to determine under what conditions the form of the function and the values of the parameters remain constant and in what way the function changes with changes in experimental conditions.

Summary

This experiment investigated the changes in the perceived spatial interval between two cutaneous stimuli when the temporal interval between them was varied. Five *O*s judged whether one or two places on the left forearm had been stimulated when the time was varied between two electrically produced stimuli separated by spatial intervals of 1.0 to 14.6 cm. The temporal interval necessary for two stimuli to be judged as occurring in two places was significantly different at different spatial intervals and for different *O*s. Curves of the form Log $T = a - bD$ were found to fit the data reasonably well. These curves imply that when the product of time and an exponential function of distance reaches a certain constant value the individual begins to discriminate between one and two.

TRANSFER OF DOUBLE ALTERNATION BEHAVIOR OF RATS IN A TEMPORAL MAZE

By HOWARD F. GALLUP, Lafayette College, and
LEONARD DIAMOND, Catholic University

The ability to double alternate in temporal mazes has been demonstrated in cats, dogs, racoons, monkeys, and in human *S*s. Though rats double alternate in spatial mazes and in modified Skinner boxes, they are incapable, according to the preponderance of evidence today, of learning to do so in temporal mazes.[1]

The interpretation of behavior in double alternation in temporal mazes, whether demonstrated or not, has usually been made in terms of symbolic processes, which are assumed to increase in complexity with the ascension in the phylogenetic scale. This is not, however, the only interpretation that may be made. Learning may be described in terms of a chain of responses—each turn being a response to the subject's proprioceptive stimuli. In the temporal maze, there is no series of places to which *S* must go; *S* must go through the maze several times en route to its goal (see Fig. 1A.). The stimulus-objects in the maze, especially those at the choice-point, cannot serve as reliable cues in a chain of responses; they can only serve to interfere with the more reliable cues from *S*'s own kinesthetic sensations. In short, an *S* learning a temporal maze must make all responses to stimuli within itself and must learn to ignore external stimuli.

Hunter concluded his work with two alternative hypotheses concerning the mode of learning a double alternation in a temporal maze: an *S* can learn to double alternate through the accumulation of internal events or through the use of symbolic processes.[2] The present authors take the accumulation of internal events to be the chaining of responses based upon kinesthetic stimulation. Schlosberg and Katz speak of symbolic activity in terms of the "cumulative effect of previous behavior, stimulus-traces, con-

* Received for publication November 13, 1958.

[1] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, rev. ed. 1954, 628; C. E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 658; and N. L. Munn, *Psychology*, 1956, 28 .9

[2] W. S. Hunter, The behavior of raccoons in a double alternation temporal maze, *J. genet. Psychol.*, 35, 1928, 374-388; W. S. Hunter and J. W. Nagge, The white rat and the double alternation temporal maze, *J. genet. Psychol.*, 39, 1931, 303-319.

tinuing brain processes, and fractional responses."[3] This definition suggests that the chaining of responses to kinesthetic cues can be considered as one kind of symbolic process, whereas Hunter defined symbolic processes as "substitute stimulus-response relationships that can be retained and later recalled to guide behavior selectively."[4] This definition seems to exclude any straightforward chaining of responses; at least it emphasizes the substitute aspect. One of the purposes of this paper was to provide a test between these hypotheses. If a chain of learned responses is broken in some way, and *S* continues to perform correctly after this break, the correct

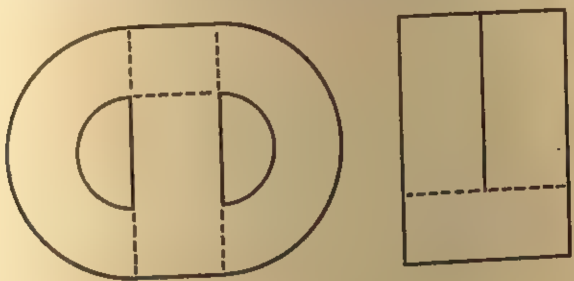


FIG. 1 DIAGRAM OF THE MAZES

A, temporal maze; B, spatial maze

performance would appear to be dependent upon substitute S-R relationships rather than upon a simple chaining of responses.

Method and procedure. The specific test was one of transfer of a double alternation from a maze requiring 360° turns to one requiring no such turns, and vice versa. Transfer was measured in terms of the number of double alternations with a limited number of training trials on the second maze in comparison to a control group which had not had the previous experience of double alternation training. A technique was devised for such training, the main features being the depletion of distinctive external stimuli to reduce interference from them as far as possible, and, in training, only one turn was learned at a time progressively until the double alternation pattern was complete.

In addition, the usual method of scoring was considered to be one which was too stringent in the sense that partial learning of the double alternation is not demonstrable. We refer to the fact that a correct double alternation has usually been considered eight consecutive turns, LLRRLLRR or RRLLRRLL. Our method, to be

[3] Harold Schlosberg and Arnold Katz, Double alternation lever-pressing in the white rat, this JOURNAL, 56, 1943, 274.

[4] Hunter, *op. cit.*, 1928, 374.

discussed below, is one which permits credits to be scored even when some turns are incorrect.

Apparatus. Two mazes were constructed (See Fig. 1), one essentially shaped like a figure '8' (Maze A), the other like a "U" (Maze B). Maze A was painted flat black; Maze B, flat blue. Both mazes were covered with heavy wire mesh which had trap doors for placing the animals into the maze from above. Both were so arranged with doors as to force an animal during training to follow *E*'s pattern in order to reach the reward (food). The doors were hung from the mesh above and could swing freely in either direction unless locked. Both mazes were 28 in. in length; Maze A was 20 in. wide, Maze B, 12 in. They were placed on the same table in the same room in which the rats were housed. Because the rats were completely familiar with the room in which they lived and were trained, they could ignore the extra-maze cues, thus reducing interference from them. Lights in the room were kept constant with an overhead fluorescent light and a 75-w. bulb suspended over the mazes.

Subjects. The *Ss*, (8 female hooded rats) were 75 days old at the beginning of the experiment. They were all from the same colony, the animal laboratory at Syracuse University. All were run on a 23-hr. deprivation-schedule on alternate days at 10 A.M. and given *ad libitum* feeding until the following day at 11 A.M. when they were deprived. Water was available at all times. A 16-day rest-period was allowed between the original and transfer trainings. The *Ss* were given 15 days of preëxperimental handling, during which they were allowed to run freely over a table top containing blocks of wood and wire mesh.

Procedure. Two *Ss* were trained in the double alternation on Maze A. One was given 20 trails on one right turn, followed by 25 trials on two right turns, 20 trials on two right turns and one left turn, and 20 trials on two right turns and two left turns. The other *S* began with a left turn and learned LLRR. The training was thus completed in four stages, each stage adding a new turn until the RRLL or LLRR pattern had been completed.[6] Two other *Ss* were used as controls. The Experimental and Control *Ss* were selected randomly and paired randomly that the inhabitants of a home cage consisted of one *S* from each group. The Control *S* was treated exactly as its experimental partner except for specific training. It was allowed to explore the same maze on the same day for the same amount of time that it took its partner to run through the daily trials. It explored under the same deprivation-schedule and was fed in the home cage immediately after being returned to it.

The same procedure was followed on Maze B with four other *Ss*. At the end of the 85 training-trials, all *Ss* were given 10 test-trials, one trial a day, under the same deprivation-schedule. A test-trial was defined as eight unrestricted turns with food on completion, regardless of correctness or incorrectness of turns.

All the *Ss* were next given double alternation training on the maze with which they had had no previous experience. Each was given 50 trials on the final stage of learning; namely, the four turns in one pattern. They were given five trials a day, the first four being training-trials and the fifth an unrestricted test-trial.

[6] See Table I for a summary of the procedure. The deviation in the number of trials at Stage 2 was necessitated by the fact that this was the hardest stage for the *Ss*. This stage required the *Ss* to pass the area where food reward had been previously received.

Scoring technique. A double alternation requires four turns: two in one direction followed by two in the opposite direction. In a given test-trial where eight turns are made, there are five sets of four consecutive turns, or as we shall label them, five 'quadruples.' For example, with turns 1 2 3 4 5 6 7 8, the animal could double alternate on turns 1234, 2345, 3456, 4567, or 5678. With a choice of either left or right on each turn, and with eight consecutive turns, there are 2^8 or 256 combinations possible. Of these 256 combinations, there are 130 in which no quadruples can be scored correct (RLRLRLRL); 95, in which one quadruple is correct (LLRR LRLR); 29, in which two quadruples are correct (LRRLLRRL); and two combinations in which three quadruples are correct (RRLLRRLL and LLRRLLRR). It

TABLE I

SUMMARY OF THE PROCEDURE

Training*	Experimental *Ss*				Control *Ss*			
	1	2	3	4	5	6	7	8
Original:								
Maze	A	A	B	B	A	A	B	B
Pattern	RRLL	LLRR	RRLL	LLRR	none	none	none	none
Transfer:								
Maze	B	B	A	A	B	B	A	A
Pattern	RRLL	LLRR	RRLL	LLRR	RRLL	LLRR	RRLL	LLRR

* *Ss* 1 and 5, 2 and 6, 3 and 7, 4 and 8 were paired.

is not possible to have four or five correct quadruples. The theoretical distribution of the combinations has a mean of 0.62 correct quadruples, with an *SD* of 0.71, when the number of trials is one. With four *Ss* in each group, and 10 test-trials per *S*, $N = 40$. The curve of expected frequencies now approximates normality, with $M = 0.62$, and $SD = 0.11$.[8]

Results. The number of correct quadruples on the test-trials for the Experimental and Control *Ss* is recorded in Table II. The totals show that the Experimental *Ss* double alternated 44 times to 4 for the Control *Ss*, the mean values being 1.1 and 0.1 correct quadruples per trial, respectively. A *t*-test of this obtained difference over the expected difference yields a value of 9.1, $p < 0.01$. We take this significant *t*-value to indicate that the Experimental *Ss* had learned something about double alternation in a temporal maze.

The number of correct quadruples on the test-trials during the transfer of training is also recorded in Table II. The original Experimental *Ss* double alternated 19 times to 1 time for the Control *Ss*, the mean values being 0.48 and 0.03 correct quadruples per trial, respectively. A *t*-test

[8] The authors are indebted to Dr. Robert Beinert of the Hobart College Mathematics Department for his help in formulating the statistical analysis.

of this difference yields a value of 4.1, which is significant at less than the 1/10% level. We regard these results as supporting the view: (1) that double alternation had been learned; and (2) that double alternation is dependent upon substitutive processes rather than upon the chaining of responses.

Maze A required a 360° turn for every trip around, and a total distance covered of about 45 in. running space. Of the 360°, 180 are accounted for by two separate right angles and the remainder by a half-circle. Thus, for a series of 8 turns, *S* had to make 16 right-angle turns, 8 half-circles, and cover a total distance of some 360 in. In Maze B, a response was scored

TABLE II

Number of Correct Quadruples after Original and Transfer Training

Original training						Transfer training					
Experimental *Ss*			Control *Ss*			Experimental *Ss*			Control *Ss*		
X	*f*	*fX*	*X*	*f*	*fX*	*X*	*f*	*fX*	*X*	*f*	*fX*
0	5	0	0	36	0	0	22	0	0	39	0
1	27	27	1	4	4	1	17	17	1	1	1
2	7	14	2	0	0	2	1	2	2	0	0
3	1	3	3	0	0	3	0	0	3	0	0
Sum = 44			Sum = 4			Sum = 19			Sum = 1		
Mean = 1.1			Mean = 0.1			Mean = 0.48			Mean = 0.03		

whenever *S* went through a door, whether or not it went to the far end of the runway. Thus one 'turn' could be made with a minimum of 10 in. of running or a maximum of 24 in. In addition, *S* could get from one runway to the other by turning in any direction, or by backing out. Thus, for a series of 8 turns, the *Ss* could each have made eight 360° turns, or none, or any combination in between, and cover a total distance of from 80 to 240 in. Observations of the *Ss* indicated a tendency to make turns and to cover the minimal distance; that is, the *Ss* went only as far as necessary, turned completely around to come out, turned around to go in, and so on. There was no 'pattern' of turning as in Maze A. Sometimes an *S* turned clockwise, sometimes not, and these were interchanged by the *Ss*. This was equally true of the *Ss* trained on Maze B and of those trained on Maze A and then transferred to B. It would appear that the pattern of responses involved in completing a series of eight consecutive turns is quite different on the two mazes. Any tendency for transfer of double alternation to take place from one maze to the other has been taken as evidence that the learning in either of these mazes is not dependent upon the

simple chaining of responses to kinesthetic cues but upon substitute S-R relationships.

It is interesting to note the mean number of correct quadruples for the Control *Ss*, as compared to the mean of the theoretical distribution, 0.62 correct quadruples, the expected value. In the first set of test-trials, the Control *Ss* displayed only four correct responses for a mean of 0.1 per trial, more than *4SD below* the expected value. This supports the belief that double alternation is a rather rare phenomenon in untrained *Ss* (rats), even with the less stringent criterion used here. In the transfer situation, where test-trials were begun earlier in training, exploratory behavior had a greater effect on the responses and the Control *Ss* performed even further below the expected value. The Experimental *Ss* also performed below the expected value on their transfer tests, but above this value on their original tests. Presumably exploratory behavior interfered here also, although the poorer showing could be attributed to the different training in the two situations or to the different number of training trials. The data do not permit a choice among these alternative hypotheses and further experimentation is under way.

Summary and conclusions. Double alternation in a temporal maze could be attributed to the chaining of responses to kinesthetic cues or to substitute S-R relationships. Four *Ss* were given training, followed by test-trials and subsequent transfer training in a new situation. The mazes used were so designed that the pattern of responses made by the *Ss* would be completely different on the two mazes. Test-trials at the end of original training indicated that the Experimental *Ss* did significantly better than the expected value and than the Control *Ss*. Transfer to the second maze also resulted in significantly better performance by the Experimental than by the Control *Ss*. The results are interpreted to show that the *Ss* (rats) can learn to double alternate in a temporal maze, and that such learning is dependent upon substitute S-R relationships rather than upon response chaining to kinesthetic cues. Correctness of a series of turns was scored in a way which permits partial correctness rather than the previously used, more stringent criterion of eight consecutive turns with no errors.

THE EFFECT OF SUBLIMINAL SHOCK UPON THE JUDGED INTENSITY OF WEAK SHOCK

By ROGER W. BLACK and WILLIAM BEVAN, Emory University

From the inception of psychophysics, it has been held that the absolute threshold provides the limiting criterion in the identification of the stimulus-correlates of the several magnitudes of sensory experience. The validity of Fechner's Law, in its general form, depends, for example, on two assumptions: that the stimulus (S) may be integrated and that it equals zero at threshold. Similarly, it relates to the limits of the reciprocity-principle—a stimulus below a critical intensity will not produce a response, regardless of its duration. Fechner's Law has, from time to time, come in for its full share of criticism on a variety of grounds, the most serious, perhaps, being that it ignores two important classes of variables—background and past experience—which influence the magnitude of the psychophysical judgment. This charge is supported by data from studies of the central-tendency effect, series-effects, interpolated stimuli, anchoring, category-limens, and adaptation-level.[1] Here still, however, one notes the persistence, at least tacitly, of the assumption that the population of inputs which give rise to a judgment is supraliminal.

Meanwhile, an experimental interest in the relation of subliminal stimuli to behavior goes back a number of years. There have been, for example, the studies of the non-randomness of response to subliminal inputs, of conditioning with subliminal stimuli, and, more recently, of perceptual defense and subception.[2] These studies have in common a preoccupation

* Received for publication April 16, 1959.

[1] Herbert Woodrow, Weight-discrimination with a varying standard, this JOURNAL, 45, 1933, 391-416; M. E. Tresselt, The influence of amount of practice upon the formation of a scale of judgment, *J. exp. Psychol.*, 37, 1947, 251-260; J. P. Guilford and D. G. Park, The effect of interpolated weight upon comparative judgments, this JOURNAL, 43, 1931, 588-599; Spaulding Rogers, The anchoring of absolute judgments, *Arch. Psychol.*, 36, 1941 (No. 261), 1-42; D. M. Johnson, *The Psychology of Thought and Judgment*, 1955, 362-367; Harry Helson, Adaptation-level as a basis for a quantitative theory of frame of reference, *Psychol. Rev.*, 55, 1948, 297-313.

[2] L. E. Baker, The influence of subliminal stimuli upon verbal behavior. *J. exp. Psychol.*, 20, 1937, 84-100; J. G. Miller, Discrimination without awareness, this JOURNAL, 52, 1939, 562-578; Baker, The pupillary response conditioned to subliminal auditory stimuli, *Psychol. Monogr.*, 50, 1938 (No. 223), 1-32; F. W. R. Taylor, The discrimination of subliminal visual stimuli, *Canad. J. Psychol.*, 7, 1953, 12-20; Elliott McGinnies, Emotionality and perceptual defense, *Psychol. Rev.*, 56, 1949, 244-251; R. C. Davis, Motor responses to auditory stimuli above and below

with the thesis that the organism responds to inputs of which it is not aware. If this proposition is accepted, one is forced to a reëxamination of the status of the threshold-assumption. By placing the problem in the psychophysical setting, one is prompted to ask not only if the organism responds to subliminal stimuli, but also if subliminal stimuli influence responses to supraliminal stimuli. Accordingly, the present experiment was intended to answer two questions. Do subliminal stimuli interpolated between supraliminal stimuli influence judgments of the latter; and, if so, is their effect similar to that associated with anchors above threshold?

The dimension chosen for study was response to shock, since an electrical stimulus can be easily controlled and because it constitutes input for which a range of readily available subliminal intensities exists. A preliminary experiment was done with seven *O*s making judgments under both control (supraliminal series alone) and experimental (supraliminal series with interpolated subliminal shocks) conditions and orders counter-balanced. While no significant-conditions main-effect was observed, two reliable interactions—shocks × conditions and *O*s × conditions—suggested the possible fruitfulness of further experimentation. Thus, the present simple two-group design was carried out.

Observers. Forty-six students in introductory psychology (20 men and 26 women) were randomly divided into two groups, the ratio of men to women in each being approximately equal. All were informed that they would be required to judge intensities of shock, but no mention of subliminal input was made. None gave any evidence of being cognizant of its use.

Apparatus. All shocks presented were square-wave pulses delivered by an Applegate Model 226 B Electronic Stimulator. Shock-duration as well as the interval between shocks, was controlled by two Hunter Model 111-C timers wired to recycle themselves. Shocks were given through two silver disks, 2 cm. in diameter, covered by saline-saturated sponge. One electrode was attached to the right wrist on the proximal side, the other on the distal side, by means of an elastic wrist-band. In addition, a 1 × 4-in. saline-moistened, sponge-covered, copper-screen electrode was attached to each ankle for the recording of changes in skin-resistance. The *GSR*s were recorded on a Brown Electronik continuous recorder with advances controlled by a third Hunter timer. Each advance was initiated 2 sec. before the onset of shock and terminated with the presentation of shock.

threshold, *J. exp. Psychol.*, 40, 1950, 107-120; R. S. Lazarus and R. A. McCleary, Autonomic discrimination without awareness: A study of subception, *Psychol. Rev.*, 58, 1951, 113-122.

Procedure. *O* was seated on a classroom chair with his back to the apparatus, the electrodes attached, and an estimate of his absolute threshold obtained by means of the method of limits with four ascending and four descending series. Shocks were 200 m.sec. in duration presented in increments of 100 μa. with a pulse-rate of 50 per sec. Next, a seven-category rating scale, varying from *very strong* through *medium* to *very weak,* was placed on the arm of *O*'s chair and instructions were provided for its use. Each *O* was reassured that the shocks would be mild, varying from a 'prick' to a 'twitch' or 'sting.' He was cautioned to make his judgments only with reference to the shocks received during the experiment and was informed that if additional categories were needed they could be added to either end of the scale. Finally, he was asked to report a judgment each time a shock was felt.

Each *O* received a series of 100 supraliminal 200-m.sec. shocks spaced at 20-sec. intervals. The series consisted of five intensities, 1500, 1800, 2100, 2400, and 2700 μa., each presented 20 times in random order. In addition, *O*s in the experimental group received a 200-m.sec. subliminal stimulus, with an intensity of 50% of the individual *O*'s threshold, at the midpoint of the 20-sec. interval between suprathreshold stimuli. In addition to the *O*'s judgments, *GSR*-deflections were recorded for all stimuli in the series, as well as for the intervals between these stimuli. Basal resistances were noted after every fourth shock. The *GSR*-apparatus was set at maximal sensitivity (change of 1000Ω for full-scale deflection, 100Ω for single-unit deflection). At the end of the session, the absolute threshold again was determined. Because of apparatus-failures, *GSR*-data were obtained only for 18 *O*s in one group and 20 in the other.

Results. The results, which were analyzed by Alexander's method,[3] are summarized graphically in Fig. 1. The fact that all points on the curve for the experimental group lie above their counterparts on the control curve suggests an affirmative answer to the first question asked in this experiment, and the reliability of this difference is demonstrated by the highly significant mean variance between groups ($p < 0.1\%$). Interpolation of imperceptible stimuli appears to enhance the judged intensity of the perceptible series-stimuli. It may be objected that, since the absolute threshold represents a range of stimulus-values from the rarely-perceptible to the always-perceptible, *O* may, on certain occasions, have been aware of the interpolated shocks and this awareness altered the judged magnitude of the

[3] H. W. Alexander, A general test for trend, *Psychol. Bull.,* 43, 1946, 533-557.

series-shocks. We are inclined to reject this possibility for several reasons. Each *O* had been instructed to make judgments of *all* shocks which he *felt*, but no judgments were ever given for the interpolated shocks. Interrogation of the *Os* after testing failed to reveal any knowledge of the interpolated stimulus. Finally, no *GSRs* could be linked with any confidence to the occurrence of these shocks. The enhancement seen in the experimental group appears to have occurred in the face of sensory adaptation. While both groups displayed an increase in threshold as a result of testing, the

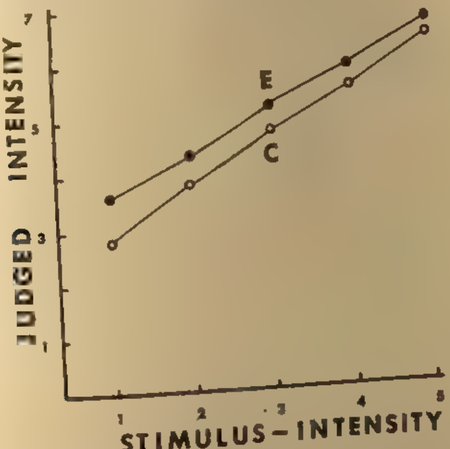


FIG. 1. AVERAGE JUDGMENTS OF INTENSITY OF FIVE SUPRALIMINAL STIMULI WITH AND WITHOUT THE INTERPOLATION OF A SUBLIMINAL ANCHOR

The physical intensities varied from 1500 to 2700 µa. in 300 µa. steps.

post-test threshold of the experimental group, which received twice the number of shocks given the control group, was slightly higher than that of the control. The present data are inconsistent with the traditional psychophysical assumption that the absolute threshold is a limiting value in the generation of psychophysical functions.

Inspection of Fig. 1 also suggests that the interpolated stimulus behaved like a supraliminal anchor introduced below the series. It drew the *IP* toward itself, but its effect was asymmetrical, being greatest upon the weaker of the series-inputs situated closer to it on the stimulus-continuum. (The variance for slopes between groups was significant at the 5% level.) Thus it would appear that the organism may incorporate subliminal along with supraliminal stimuli in the derivation of subjective norms for judg-

ment, and the criterion of perceptibility cannot be applied in defining the properties of input relevant for pooling.[4]

Lastly, with the *GSR*-apparatus set at maximal sensitivity, deflections were observed in 92.2% of the 10-sec. intervals following series-stimuli, in only 9.5% of the 10-sec. intervals following interpolated shock in the experimental group, and in 9% of the corresponding 'empty' intervals for the control group. The difference between frequency of *GSR*s observed in the supraliminal and subliminal or empty intervals is highly reliable. Chi-squares computed for individual *O*s all were highly significant, the smallest being 37.2. Meanwhile, the average frequency of response by the experimental group during the subliminal period was not reliably in excess of that seen in the controls during the 'empty' period ($U = 176$, $m/n = 20/18$, $p > 0.10$).

It would thus appear that *GSR*s during the subliminal interval were 'spontaneous' and not systematically related to the subthreshold shock. This failure for the subliminal shock to produce a *GSR*, although it produced a significant change in judgment, points up two things. First, this is another instance in which a behavioral measure appears to be more sensitive than a widely used physiological measure as an index of the influence of a physical variable. Secondly, the notion of arousal, so closely linked, at least since the time of Cannon, to the autonomic nervous system, warrants more definitive formulation. It may not be inappropriate to suggest relating the present findings to the function of the thalamic portion of the reticular activating system, while those from the studies of perceptual defense may be thought to reflect the activity of the basal diencephalic and midbrain parts of this same complex.[5]

Summary. The present experiment investigated the effect of subliminal electrical shocks upon the judged intensity of supraliminal shocks. Two groups of *O*s received 100 shocks above threshold, 20 at each of five in-

[4] The obtained *IP* for the control group was 1841 μa., for the experimental group, 1671 μa. Three estimates of the subjective norm were computed to test the adequacy of the quantitative aspect of the theory of adaptation-level: the arithmetic mean; the logarithmetic mean with Helson's correction (+0.75d) applied for interval-size; and the exponential or power-mean. The corrected log mean and the power-mean predicted with approximately the same efficiency, but only moderately well (error = 15%). Corrected log mean: $C = 2281$, $E = 1486$; power mean: $C = 2181$, $E = 1458$. Since, at this point, we are not preoccupied with precise quantitative prediction, no attempt has been made to determine constants for purpose of correction. Meanwhile, it would appear that the subliminal anchor is less potent than a supraliminal counterpart might be.

[5] Herbert Jasper, Reticular-cortical systems and theories of the integrative action of the brain, in H. F. Harlow and C. N. Woolsey (eds.) *Biological and Biochemical Bases of Behavior*, 1958, 37-63.

tensities, to be rated on a seven-point scale. One group, in addition, received shocks of 50% threshold-intensity interpolated between the series-stimuli. *GSR*-deflections were recorded for both supra- and subliminal presentations.

The mean intensity of the series-judgments for the group receiving the interpolated shock was significantly greater than for the control group. Further, the interpolated shock behaved like an anchor below the series, effecting the greatest enhancement for the series-stimuli closest to it on the stimulus-continuum. While *GSR*-deflections occurred consistently with the presentation of the series-stimuli, none could be confidently related to the presentation of the subliminal stimuli. The implications of these data for pooling theory and for the concept of arousal were noted.

DEVELOPMENTAL AND INTELLECTUAL PROCESSES IN SIZE-DISTANCE JUDGMENT

By NOËL JENKIN, University of Sydney, and SALLY M. FEALLOCK, University of Virginia Hospital

In a study by Beyrl of children and adults, evidence was found to support the view that the tendency toward size-constancy develops progressively with age.[1] Frank subsequently argued that this result was an artifact of the experimental situation, and that constancy of apparent visual size actually was as well developed in the 2-yr.-old as in the 10-yr.-old child.[2] Burzlaff also found that evidence favoring a developmental view of size-constancy depended on the nature of the experimental procedure employed.[3] More recent work has very consistently indicated that differences in performance between children and adults are not due to any artifact but are clearly related to development.[4]

The present study therefore seeks the answer to a hitherto unresolved question: Is the correlation between age and judgment of size-distance mediated by the growth of intelligence, or is it a developmental relationship independent of intelligence? The existing literature relating to this problem is sparse and contradictory. Locke presented a case for the view that intelligence and degree of perceptual constancy are inversely related.[5] Thouless intercorrelated measures of distance-judgment with test-scores

* Received for publication May 18, 1959. This study was conducted when the authors were at The Training School at Vineland, N.J. Thanks are due to the Superintendent and staff of the Vineland State School, whose coöperation made possible the collection of some of the data.

[1] Franz Beyrl, Über die Grössenauffassung bei Kindern, *Z. Psychol.*, 100, 1926, 344-371.

[2] Helene Frank, Untersuchung über Sehgrössenkonstanz bei Kindern, *Psychol. Forsch.*, 7, 1926, 137-145; Die Sehgrössenkonstanz bei Kindern, *ibid.*, 10, 1927, 102-106.

[3] Wilhelm Burzlaff, Methodologische Beiträge zum Problem der Farbenkonstanz, *Z. Psychol.*, 119, 1931, 177-235.

[4] Marc Lambercier, La constance des grandeurs en comparaisons sériales, *Arch. Psychol., Genève*, 31, 1946, 79-282; La configuration en profondeur dans la constance des grandeurs, *ibid.*, 31, 1946, 287-324; Jean Piaget and Marc Lambercier, La comparaison visuelle des hauteurs à distances variables dans le plan fronto-parallele, *ibid.*, 29, 1943, 175-253; La comparaison des grandeurs projectives chez l'enfant et chez l'adulte, *ibid.*, 33, 1951, 81-130; Grandeurs projectives et grandeurs réelles avec étalon éloigné, *ibid.*, 35, 1956, 257-280; H. P. Zeigler and Herschel Leibowitz, Apparent visual size as a function of distance for children and adults, this JOURNAL, 70, 1957, 106-109.

[5] N. M. Locke, Perception and intelligence: Their phylogenetic relation, *Psychol. Rev.*, 45, 1938, 335-345.

of intelligence and found negative coefficients.[6] These results, while given limited support by Sheehan,[7] seem inconsistent with the repeated finding that the tendency toward size-constancy is positively related to age. A need clearly exists for a study in which the developmental trend is examined with the factor of intelligence under adequate control. The present experiment was planned to satisfy this requirement by comparing the judgments of normal children from two age-groups with the judgments of mentally retarded persons who had no symptoms or history of damage to the nervous system. To separate the role of intelligence from that of other developmental factors, it was intended that the retarded group should approximate the chronological age of the older normal group and the mental age of the younger normal group. While not crucial to the purpose of the experiment, it was decided also to add an adult group to clarify further the developmental picture. Two main findings were expected. The first, based on the weight of existing evidence, was a developmental trend in the direction of increasing size-estimates with increasing age. The second arises from the *a priori* view that judgments of size-distance lack the essential properties which give rise to individual differences in intelligence test-scores. It was anticipated, therefore, that the performance of the retarded group would resemble that of the older rather than the younger normals.

Two additional variables were included in the design of the experiment. The first, angular separation between the stimulus-objects, was based on Frank's claim that such differences account for the variation in performance between younger and older children.[8] The second, distance of the comparison-object, was incorporated because of Jenkin's suggestion that increases of apparent size with increasing distance are related to visual learning.[9]

Subjects. Four groups of *Ss* were used, each group numbering 20 cases and including both sexes. Cases of gross visual defect were excluded from each group. The group of young children (hereafter referred to as 'children') consisted of normally achieving pupils from a Vineland grade school. The median chronological age of this group was 8 yr. 3 mo. The group of normal adolescents (hereafter

---

[6] R. H. Thouless, Individual differences in phenomenal regression, *Brit. J. Psychol.*, 22, 1932, 216-241.

[7] M. R. Sheehan, A study of individual consistency in phenomenal constancy, *Arch. Psychol.*, 31, 1938, (No. 222), 1-95.

[8] Frank, *op. cit.*, 1927, 102; Kurt Koffka, *Principles of Gestalt Psychology*, 1935, 238-240.

[9] Noël Jenkin, Effects of varied distance on short-range size-judgments, *J. exp. Psychol.*, 54, 1957, 327-331; A relationship between increments of distance and estimates of objective size, this JOURNAL, 72, 1959, 345-363.

referred to as 'adolescents') consisted of normally achieving students from two local junior high schools. The median chronological age was 13 yr. 7 mo. The 'retarded' group consisted of residents of The Training School at Vineland and the Vineland State School. The medical history and present status of each member of this group were free from any indication of damage to the nervous system. Most individuals in this group were classified in the records of the respective institutions as of familial etiology. The median chronological age was 15 yr. 10 mo. and the median mental age was 8 yr. 2 mo. The 'normal adult' group consisted mostly of employees of The Training School at Vineland, together with a few persons associated indirectly with the school. The median chronological age of this group was 26 yr. 6 mo.

Apparatus. The apparatus consisted of white cardboard squares. The standard card, 4 in. on a side, was attached by a steel wire to a ringstand and was placed directly in front of *S* at a distance of 320 in. The comparison-series was a set of squares ranging from 2.75-6.00 in. by steps of ⅛ in. They were individually displayed at the same height as the standard upon a platform which was situated on *S*'s right. Concealed pins and a slight depression, invisible to *S*, served to maintain the comparison-cards in vertical position in the middle of the platform. Homogeneous fields against which both stimuli were viewed were provided by wooden screens painted a dull green. Little daylight entered the room, and the stimulus-objects were illuminated by 75-w. floodlights so placed as to give an illumination of 11 ft.-C. (measured by a Macbeth illuminometer) from a correctly positioned square.

Procedure. Angular separation was varied by placing the comparison-object either 19° or 87° from *S*'s line of vision when viewing the standard object. At both angles, the distance of the comparison-object remained constant at 80 in. from the point of observation. Variation of distance was accomplished by placing the comparison-object either 20 in. or 160 in. from the point of observation. When the distances were varied, separation between standard and comparison remained constant at 87°. Since all the *S*s experienced all the experimental conditions, a subordinate variable was that of order. Each main group was divided into two subgroups, A and B. Subgroup A experienced the order: 19°, 87°, 160 in., 20 in. Subgroup B experienced the order 20 in., 160 in., 87°, 19°. The method of limits was used; half of each subgroup began with an ascending trial and half with a descending trial.

The height of each stimulus-object was matched with *S*'s eye-level prior to each session. Instructions to *S* emphasized a judgment of physical equality. Two practice-trials were conducted for each *S*, and an equal amount of practice under the four experimental conditions was uniformly distributed throughout each main group.

Results. Table I shows the mean size-matches of the four groups. The first question to be asked is whether or not the data confirm previous studies which have shown a progression toward or beyond size-constancy with increasing age. Table I indicates that the means do in fact lie in the expected relationships, the three normal groups making mean matches which become larger as the median ages increase. An analysis of variance

of the normal groups across all conditions indicated a highly significant difference for the 'Groups' variable ($F = 12.48$, $p < 0.001$, with 2 and 57 *df.*).

To study the significance of differences between pairs of groups, including the retarded cases, separate analyses of variance were computed for each set of experimental conditions. Between-group comparisons were based on 1 and 36 *df.* in each case. Under the varied-distance conditions, children differed significantly from adolescents ($F = 6.63$, $p < 0.02$), and adolescents differed significantly from adults ($F = 7.11$, $p < 0.02$). Under the angular-separation conditions children differed significantly from adolescents ($F = 4.85$, $p < 0.05$), but adolescents did not differ significantly from adults ($F = 3.80$, $p < 0.1$).

TABLE I

AGE-STATISTICS AND MEAN SIZE-MATCHES OF FOUR GROUPS

(Ages are expressed in years and months, size-matches in inches)

		Children	Adolescents	Retarded*	Adult
Ages	Median	8·3	13·7	15·10	23 6
	Mean	8·2	13·1	16·3	25·5
	SD	1·3	1·2	3·4	6·6
Mean size-matches	19°	3.74	4.09	4.13	4.41
	87°	3.81	4.12	4.18	4.45
	20 in.	3.83	4.21	4.14	4.65
	160 in.	3.81	4.05	3.96	4.36
	Grand mean	3.80	4.12	4.10	4.47

* The retarded group had a median mental age of 8 yr. 2 mo.

No significant differences were found between the adolescent and the retarded groups. Comparison of the retarded cases with the children, however, yielded a significant difference ($F = 6.19$, $p < 0.05$) under the varied-distance conditions and another significant difference ($F = 7.14$, $p < 0.02$) under the angular-separation conditions. This pattern of results confirms the main hypotheses of the study.

Interactions between the various pairs of groups, on the one hand, and the conditions of angular separation, on the other hand, were all nonsignificant ($p > 0.05$). Differences between the angular-separation conditions across groups also were nonsignificant. Variability between Distances across Groups was significant in all comparisons where the adult group was involved (Children *vs.* Adult, $F = 4.65$, $p < 0.05$; Retarded *vs.* Adult, $F = 4.60$, $p < 0.05$; Adolescent *vs.* Adult, $F = 11.75$, $p < 0.01$; *df.* = 1 and 36 in each comparison) and nonsignificant between pairs of groups

which did not include adults. Table I indicates that the direction of the differences was toward smaller matches as the comparison-object was placed further from *S*, a result which confirms Jenkin's earlier findings. All analyses in the present study were evaluated by means of two-tailed tests.

Discussion. The results of this experiment unequivocally confirm the two main hypotheses. There is a clear progression with increasing age toward and beyond size-constancy. There is also clear evidence that this progression with age is not mediated by the kinds of intellectual ability which are measured in a typical intelligence test.

As Zuckerman and Rock have pointed out, the interaction of retinal size and perceived distance may be learned or innate, but the same is true of distance cues.[10] The present experimenters and their predecessors have obtained data which are consistent with the view that *either* the interaction of retinal size and perceived distance *or* distance-cues are a product of learning. In the light of animal experiments,[11] and studies of pre-verbal children,[12] it seems most meaningful at present to suppose, as Gestalt writers have suggested, that the process of interaction is given by an innate organization. In this case, results such as obtained in the present study are best interpreted as indicating a developing appreciation of depth.

In the present study, the varied conditions of distance gave results which are consistent with the above hypothesis. Transfer of the position at which the comparison-object was presented (involving only a short interval in depth) had little effect on the mean size-judgments of the children, whereas, with the adolescents and especially the adults, a significant difference occurred. The fact that size-matches of children are less sensitive to changes in depth than are those of adults is a corollary to the main results, and is also consistent with the view that, as age increases, depth-cues play a more important part in size-distance judgment.

Summary. Objective, binocular size-matches under conditions of full illumination were obtained from three groups of normal persons (children, adolescents, and adults) and from one group of mentally retarded persons who had neither past history nor present symptoms of damage to the nervous system. The children made a mean size-match smaller than the value

---

[10] C. B. Zuckerman and Irvin Rock, A reappraisal of the roles of past experience and innate organizing processes in visual perception, *Psychol. Bull.*, 54, 1957, 269-296.

[11] Locke, *op. cit.*, 335-345.

[12] R. M. Cruikshank, The development of visual size constancy in early infancy, *J. genet. Psychol.*, 58, 1941, 327-351; Frank, *op. cit.*, 1927, 102.

required for size-constancy. The adolescents slightly exceeded size- constancy, and the adult group exceeded this value still further. Significant differences existed between the mean matches of the three groups. The retarded group, which had a median mental age similar to that of the children but a chronological age which approximated that of the adolescents, made size-matches significantly different from those of the children but similar to those of the adolescents. Evidence is thus provided for the view that there is a relationship between increasing age and larger size-estimates, and that this progression with age is not mediated by developing intelligence. Additional findings indicate that children are less susceptible than adults to variations in depth, and that varying the angular separation between standard and comparison-object does not have any significant effect on size-matches.

DEGREE OF POLARIZATION AND SCORES ON THE PRINCIPAL FACTORS FOR CONCEPTS IN THE SEMANTIC ATLAS STUDY

By JAMES J. JENKINS, University of Minnesota

In earlier publications Jenkins, Russell, and Suci made available an atlas of semantic profiles for 360 concepts rated on 20 bipolar scales and subsequently announced the availability of a table of distances in the 20-scale space between all possible pairs of the atlas concepts.[1] To facilitate further use of these data, two additional kinds of information have recently been calculated: (1) the degree of polarization of each concept; and (2) the score of each concept on the three most salient dimensions of meaning isolated in earlier studies.

(1) Degree of polarization. Degree of polarization is defined as the extent to which the profile for a given concept deviates from a completely neutral profile; that is, from a profile which has a value of four for each of the 20 seven-step scales. This is equivalent to measuring how far each 'concept point' is from the point of origin of the semantic space where the point of origin is considered to be the point which signifies 'meaninglessness.'[2] In a loose sense this distance might be called the 'meaningfulness' of the concept. Degree of polarization appears, however, to describe the measurement more accurately and is less likely to be confused with other terms which entail different operations (such as Noble's 'meaningfulness').[3] To be sure, 'degree of polarization' is related to other constructs referring to 'meaning' and 'meaningfulness,' but it is by no means identical with them.

Degree of polarization has been calculated by the following formula:

---

* Received for publication May 28, 1959. The data which made this study possible were gathered under the Office of Naval Research Contract, N8 onr 66216. The writer is grateful to the Social Science Research Council for a Faculty Research Fellowship during which part of the study was performed, and to the graduate school of the University of Minnesota for a grant to aid the research.

[1] J. J. Jenkins, W. A. Russell, and G. J. Suci, An atlas of semantic profiles for 360 words, this JOURNAL, 71, 1958, 688-699; A table of distances for the semantic atlas, this JOURNAL, 72, 1959, 623-625.

[2] See C. E. Osgood, G. J. Suci, and P. H. Tannenbaum, *The Measurement of Meaning*, 1957, 96.

[3] C. E. Noble, An analysis of meaning, *Psychol. Rev.*, 59, 1952, 421-430.

$$D_4 = \sqrt{\sum_1^{20}(x-4)^2}$$

where x is the value of the concept on a particular scale.

The possible range of D_4 is from zero to 13.4, with 20 seven-step scales. Small values of D_4 indicate that the mean profile for the concept has little deviation from neutral values either because none of the *Ss* has any clear connotative meaning for the concept (*e.g.* GOJEY–2.67; BODKIN–2.30; LEMUR–2.24) or because there is little agreement among the *Ss* as to the connotative meaning. (This is presumably the case with DREAM–2.96; NEGRO–2.73 and SOCIALISM–2.78.) High values of D_4 indicate concepts whose profiles deviate markedly from the midpoint. Such concepts have strong connotations which are rated consistently by most of the *Ss* (*e.g.* GOD–8.35; DOCTOR–8.15; TORNADO–8.23).

TABLE I

SCALES USED IN COMPUTING FACTOR SCORES

Scales	Factor loadings		
	I	II	III
Factor I:			
Good-Bad	1.00	.00	.00
Timely-Untimely	.37	.04	.04
Kind-Cruel	.52	−.28	.00
Beautiful-Ugly	.52	−.29	−.02
Successful-Unsuccessful	.51	.08	.29
Important-Unimportant	.38	.04	.31
True-False	.50	−.03	.01
Wise-Foolish	.57	.06	.11
Factor II:			
Hard-Soft	−.24	.97	.00
Masculine-Feminine	−.14	.47	.03
Strong-Weak	.30	.40	.10
Factor III:			
Active-Passive	.17	.12	.98
Fast-Slow	.01	.26	.35
Excitable-Calm	−.15	.03	.26

Research by Dicken on semantic generalization across sets of concepts which have similar connotative meaning suggests that the degree of generalization is in part dependent on the degree of polarization of the concepts.[4] Over several studies in generalization he obtained a correlation between the frequency of the generalized response and the degree of polarization of +0.66.

[4] C. F. Dicken, Connotative meaning as a determinant of stimulus-generalization, Doctoral dissertation, University of Minnesota, 1957.

TABLE II

DEGREE OF POLARITY AND FACTOR SCORES FOR THE SEMANTIC ATLAS

Words	D_4	Factor		
		I	II	III
1. ABORTION	6.81	5.52	4.41	5.08
2. ABRUPT	5.06	4.46	2.60	5.83
3. ADORNED	4.68	3.07	4.87	4.01
4. AFRAID	4.81	4.95	4.49	5.26
5. AGILE	5.18	2.97	2.92	5.60
6. AMERICA	7.44	2.04	3.15	5.35
7. ANGER	6.36	4.98	2.73	6.23
8. ARCHED	3.62	3.67	3.23	4.69
9. ARGON	2.98	3.41	3.50	4.48
10. ARMY	5.60	3.71	1.98	5.11
11. ART	5.14	2.65	4.03	3.94
12. BABY	7.72	2.47	5.77	4.51
13. BAD	6.01	5.51	3.90	4.98
14. BARN	5.08	2.80	2.81	3.46
15. BASE	3.99	3.70	2.46	4.20
16. BATH	6.91	2.16	4.77	3.18
17. BEAUTIFUL	7.44	2.14	5.22	4.07
18. BED	5.95	2.38	4.47	2.91
19. BEGGAR	6.02	5.18	4.49	2.69
20. BIBLE	6.66	2.11	3.05	4.16
21. BIRTH	7.40	1.99	4.53	4.86
22. BIRTH CONTROL	4.12	2.86	3.68	4.52
23. BITTER	4.96	5.02	2.98	4.92
24. BLEAK	4.16	4.72	3.88	2.98
25. BLOCK	5.18	3.90	2.12	3.88
26. BLUE	5.58	2.71	4.58	3.29
27. BOAT	5.12	2.80	2.63	4.64
28. BODICE	6.59	2.73	5.73	4.13
29. BODKIN	2.30	4.40	3.58	3.60
30. BOTTOM	3.48	4.49	3.72	3.19
31. BOULDER	5.96	4.06	1.68	2.96
32. BOX	4.40	3.38	2.41	3.96
33. BOY	6.19	2.96	2.33	5.55
34. BRAVE	7.25	2.43	1.98	5.00
35. BREAD	6.60	2.46	4.35	3.25
36. BRIGHT	6.18	2.50	3.02	5.30
37. BRILLIANT	6.14	2.32	3.05	4.97
38. BRINY	3.42	4.15	2.78	4.53
39. BRISTLY	4.63	4.07	2.16	5.19
40. BROKEN	4.59	5.00	3.81	4.68
41. BROTHER	6.49	2.49	2.29	4.95
42. BUTTER	4.85	2.99	4.64	3.07
43. CALM	6.75	2.22	4.31	1.97
44. CANDY	5.08	3.19	4.75	3.81
45. CAP. PUNISH.	5.85	4.77	2.30	5.33
46. CAR	6.93	2.64	2.26	5.58
47. CHAIR	4.45	3.04	3.66	2.73
48. CHARMING	6.11	2.44	5.16	4.15
49. CHILD	6.53	2.69	4.56	5.87
50. CHURCH	6.51	2.14	3.59	4.04
51. CITY	5.84	3.32	2.70	6.02
52. CLEAN	6.75	2.01	3.89	4.29
53. CLUMSY	4.94	5.19	3.90	3.43
54. COAL	5.22	3.13	2.29	4.37
55. COARSE	4.59	4.78	2.31	4.51
56. COILED	4.41	4.13	3.40	5.22
57. COLD	5.27	4.43	2.72	4.48
58. COLOR	5.85	2.61	4.23	4.58
59. COMFORT	6.24	2.40	5.00	2.70
60. COMPLETE	6.62	2.06	2.77	3.94
61. CONSC. OBJ.	2.81	4.14	3.57	3.35
62. CONSTANT	5.22	2.72	2.72	3.53
63. CONTINUOUS	3.21	3.28	3.33	3.99
64. CONTROVERSY	4.66	3.51	2.93	5.71
65. COP	6.56	2.74	1.87	4.83
66. COURAGE	6.79	2.24	2.48	4.65
67. CRIMINAL	7.02	5.55	2.82	5.65
68. CROOKED	5.75	5.34	3.50	4.78
69. CRUDE	4.47	5.11	3.17	4.03
70. CURLED	4.49	3.55	5.13	3.78
71. CURVED	5.31	3.10	4.95	3.94
72. CUSHION	6.83	3.00	5.97	2.29
73. DANGER	5.50	4.62	2.70	5.90
74. DARK	2.91	3.83	3.29	3.36
75. DAWN	7.51	2.09	4.60	3.56
76. DEBATE	5.40	3.03	2.61	5.82
77. DEEP	3.04	3.78	2.86	3.56
78. DEFORMED	5.99	4.96	3.87	2.96
79. DELAYED	4.54	4.98	4.68	3.11
80. DELIBERATE	4.74	3.13	2.52	4.10
81. DEVIL	7.40	5.22	2.06	5.85
82. DIM	4.04	4.08	4.82	2.47
83. DIRT	3.96	4.64	3.17	4.11
84. DISCOMFORT	5.30	5.05	3.30	4.57
85. DISCORDANT	4.66	5.10	3.32	5.10
86. DITCH-DIGGER	5.22	4.42	1.84	4.51
87. DIVORCE	6.28	5.29	3.83	5.32
88. DOCTOR	8.15	1.87	2.57	4.47
89. DOOR	5.81	3.16	2.25	4.05
90. DOUGH	4.83	3.44	5.30	3.18
91. DOWNY	4.83	3.48	5.73	2.94
92. DREAM	2.96	3.88	4.23	4.56
93. DREARY	6.25	5.27	4.66	2.16
94. DUSKY	4.17	3.42	4.66	2.38
95. EASY	4.23	3.28	5.37	3.14
96. EAT	6.41	2.49	4.01	4.45
97. EATING	7.26	2.18	3.60	4.60
98 EDGED	4.38	4.04	2.60	5.22
99. EFFORT	5.97	2.48	2.56	4.97
100. ELEGANT	3.96	3.24	4.32	3.92
101. ELEVATED	3.82	3.12	3.09	4.38
102. ENGINE	7.29	2.76	1.72	6.07
103. EVEN	3.71	3.13	3.87	3.20
104. FAITH	7.69	1.70	3.77	3.70
105. FAMILY	7.06	1.94	3.73	4.42
106. FAR	2.71	3.58	3.25	3.93
107. FARM	6.29	2.40	2.60	3.78
108. FAST	6.21	3.30	2.39	6.31
109. FAT	5.51	4.51	5.37	2.91
110. FATHER	6.02	2.53	2.45	4.71
111. FEAR	4.64	4.83	3.89	5.31
112. FEATHER	5.59	3.25	6.13	3.21
113. FERVID	3.33	3.74	3.38	5.42
114. FEVERISH	4.01	4.68	3.90	5.40
115. FIERY	5.67	4.32	3.01	6.31
116. FIFTH AMEND.	3.12	3.40	3.41	4.26
117. FIRE	6.76	3.85	2.21	6.38
118. FIRM	5.65	2.77	2.01	3.91
119. FLAMING	6.10	3.74	3.23	6.42
120. FLASHY	5.22	4.59	3.23	6.08
121. FLEA	4.92	4.84	3.97	5.50
122. FLEECY	5.01	3.23	5.56	3.23
123. FLOWER	7.49	2.30	5.70	3.01
124. FLOWERS	7.29	2.38	5.62	3.06
125. FOOD	6.42	2.25	3.87	4.36
126. FOOT	4.21	3.32	3.55	4.99
127. FOREIGNER	3.07	3.48	3.34	4.30
128. FRAGRANT	5.73	2.85	5.48	3.86
129. FRAUD	5.69	5.45	3.31	4.85
130. FRIGHTFUL	5.55	5.01	3.56	5.72
131. FRIGID	4.49	4.77	3.79	3.09
132. FROSTY	3.73	3.63	3.22	4.59
133. FULL	3.51	3.09	3.92	3.49
134. GARMENT	5.80	2.67	4.79	3.32
135. GIRL	6.51	2.74	5.96	4.46
136. GLARING	5.21	4.93	2.88	5.22
137. GLEAMING	5.17	2.94	3.38	5.05
138. GLOOMY	5.14	5.09	4.08	2.75
139. GLOVE	4.55	2.93	4.82	3.38
140. GLOW	5.13	2.84	5.06	3.86

TABLE II (*Continued*)

Words	D_4	Factor I	Factor II	Factor III
141. GLOWING	4.21	3.02	4.92	3.81
142. GOD	8.35	1.68	2.41	4.19
143. GOJEY	2.67	4.58	3.81	4.15
144. GRACEFUL	7.09	2.17	5.18	3.51
145. GRADUAL	4.94	3.06	4.31	2.36
146. GREEN	3.92	3.18	4.21	3.84
147. GRIEF	4.94	4.89	3.84	4.26
148. HAMMER	6.28	3.29	1.60	5.52
149. HAND	5.53	2.55	2.39	4.89
150. HAPPY	7.85	1.84	4.05	5.11
151. HARD	5.03	4.09	1.94	4.33
152. HARMONIOUS	6.44	2.24	4.02	3.85
153. HASTY	4.37	4.43	3.71	6.20
154. HATE	6.96	5.71	3.22	5.29
155. HEAL	6.05	2.17	3.86	4.07
156. HEALTH	5.59	2.41	3.20	4.30
157. HEARTLESS	6.15	5.56	3.02	4.21
158. HEAVY	4.33	4.20	2.00	3.27
159. HIGH	4.23	3.10	2.80	4.54
160. HIT	6.06	4.68	2.00	6.21
161. HOLY	7.34	1.88	3.99	3.44
162. HOME	7.24	1.90	3.97	4.13
163. HOSPITAL	6.17	2.24	3.62	4.34
164. HOT	3.94	3.75	3.61	5.69
165. HOUSE	6.60	2.13	3.58	3.47
166. HUNGRY	3.73	4.23	3.83	4.69
167. HURRIED	4.67	4.03	3.40	6.39
168. HURT	5.48	5.08	3.04	5.37
169. INCOME	6.25	2.37	2.66	4.90
170. INDIFFERENT	5.32	5.19	4.28	2.44
171. INFERIOR	5.28	5.26	4.69	3.14
172. INSANE MAN	6.07	5.24	3.69	4.77
173. INTERMITTENT	6.37	4.28	3.98	4.17
174. JELLY	4.97	3.43	5.55	3.68
175. JEW	4.54	3.19	2.92	5.17
176. JOY	7.54	1.98	3.91	5.63
177. JUMP	5.12	3.69	2.50	6.30
178. JUSTICE	6.79	2.20	2.58	4.31
179. KITCHEN	7.23	2.20	4.60	5.07
180. KITTENS	6.19	3.35	6.04	4.70
181. KNIFE	5.18	4.20	1.96	5.30
182. LADY	6.50	2.45	5.71	4.08
183. LAGGING	5.74	5.07	4.92	2.13
184. LAKE	6.17	2.71	4.13	3.12
185. LAMP	5.50	2.46	3.83	3.22
186. LATE	5.83	5.19	4.40	3.32
187. LAZY	6.86	5.12	5.06	1.66
188. LEADERSHIP	7.01	2.32	2.30	4.91
189. LEG	5.55	2.84	4.13	5.25
190. LEISURELY	5.07	3.07	4.85	2.04
191. LEMON	5.27	3.43	3.27	4.90
192. LEMUR	2.24	3.78	4.58	3.94
193. LENIENT	4.64	2.93	5.05	2.99
194. LEPER	4.67	4.65	4.10	3.52
195. LIFT	4.96	3.00	2.35	5.01
196. LIGHT	7.05	2.03	3.70	5.05
197. LINGERING	4.01	4.22	4.71	2.48
198. LIQUOR	5.19	4.69	2.46	5.50
199. LIZARD	3.16	4.64	3.94	3.85
200. LOFTY	2.87	3.38	3.84	3.12
201. LONG	3.92	3.85	2.89	3.03
202. LOVEABLE	7.12	2.11	4.98	4.03
203. LOVELY	6.15	2.50	5.52	3.55
204. LOW	3.18	4.67	4.11	3.39
205. LUMINOUS	3.76	3.04	4.42	4.14
206. LUSCIOUS	5.30	2.95	4.74	4.52
207. MAD	6.38	5.29	2.92	6.15
208. MAJOR-OPINION	3.68	3.65	3.05	5.04
209. MALLET	5.15	3.97	1.71	4.96
210. MAN	5.90	2.81	2.40	4.96
211. ME	3.65	3.04	4.03	4.69
212. MEMORY	5.21	2.62	3.79	4.23
213. MILD	5.76	2.98	5.43	2.16
214. MILLIONAIRE	5.95	2.83	2.32	4.83
215. MIND	6.32	2.36	3.11	5.10
216. MINISTER	6.21	2.38	3 25	4.19
217. MOLD	3.30	3.82	4.07	2.99
218 MONEY	5.78	2.70	2.35	5.32
219. MOON	6.63	2.57	4.27	2.36
220. MOSQUITO	5.05	4.82	3.94	5.74
221. MOTHER	6.89	2.10	5.03	4.56
222. MOUNTAIN	6.53	2.79	1.78	3.38
223. MUSIC	7.67	2.00	4.27	4.66
224. NAIL	5.82	3.42	2.03	4.00
225. NARROW	4.17	4.56	3.78	3.79
226. NASTY	5.87	5.39	3.51	5.47
227. NEGRO	2.73	3.77	3.63	3.90
228. NEUROTIC MAN	5.60	5.36	3.98	4.78
229. NICE	6.78	1.98	4.25	3.81
230. NURSE	7.84	1.96	4.88	3.34
231. OBSCURE	3.35	4.60	3.92	3.10
232. OBVIOUS	2.87	3.31	3.46	4.28
233. OCEAN	6.57	3.00	2.46	5.64
234. OVERCAST	3.79	4.76	3.65	3.51
235. PAIN	4.98	4.73	3.16	5.33
236. PATRIOT	6.57	2.46	2.41	5.36
237. PEACE	7.33	1.89	4.30	2.62
238. PIANO	5.71	2.36	3.51	4.62
239. PIG	5.00	4.04	4.40	3.40
240. PIGMENT	3.27	3.29	4.08	4.17
241. PLAIN	3.91	3.97	3.77	2.89
242. PLIABLE	3.62	3.75	5.00	3.43
243. POLITICIAN	5.25	3.46	2.44	5.58
244. PRETTY	6.99	2.49	5.59	4.30
245. PROGRESS	7.04	2.16	2.40	5.09
246. PUNGENT	3.49	4.28	2.87	4.98
247. PUPPIES	7.00	3.06	5.35	5.60
248. PUTRID	5.46	5.56	3.45	4.72
249. PYTHON	4.65	4.27	2.91	4.63
250. QUOTA	3.41	3.57	2.90	4.58
251. RADIANT	6.88	2.22	4.25	5.40
252. RAGE	6.97	5.48	3.48	6.49
253. RANCID	4.36	4.95	3.49	4.10
254. RAPID	6.49	3.23	2.50	6.59
255. RED	5.46	3.67	3.08	5.96
256. REFINED	5.15	2.52	4.26	3.35
257. RELAXED	6.67	2.48	4.59	1.80
258. RELIGION	6.40	2.19	3.42	3.84
259. RICH	4.87	2.86	3.21	4.97
260. RIGHT	6.65	2.02	3.21	4.32
261. RIGID	5.15	4.24	2.11	3.83
262. RIPE	6.39	2.43	4.72	3.74
263. RIVER	5.72	2.87	3.20	5.09
264. ROMPIN	2.58	3.98	3.86	5.21
265. ROOT	5.11	3.10	2.69	3.28
266. ROSE	7.08	2.47	5.62	3.23
267. ROUGH	5.17	4.86	2.09	5.22
268. ROUND	5.22	3.31	5.10	3.48
269. RUGGED	5.35	3.70	1.73	4.99
270. RUN	4.96	3.88	2.77	6.14
271. RUSSIAN	4.62	3.82	2.04	4. 6
272. SAVORY	5.22	2.83	3.98	4.78
273. SCALDING	5.99	5.00	3.16	5.83
274. SCENE	4.94	2.71	4.18	3.64
275. SCIENTIST	6.66	2.26	2.62	4.23
276. SCORCHING	5.49	5.08	3.35	5.49
277. SEE	5.83	2.38	3.52	4.80
278. SEVERE	4.80	4.57	2.49	4.76
279. SEX	7.29	2.20	4.19	5.65
280. SHADY	3.38	3.56	4.47	2.78
281. SHINY	5.43	2.85	2.91	5.06
282. SHORT	2.31	3.68	4.61	3.49
283. SICKNESS	5.78	5.06	3.49	3.28
284. SILK	6.17	2.84	5.67	3.18

TABLE II (*Continued*)

Words	D_4	Factor I	II	III
285. SIN	5.87	5.08	3.64	5.08
286. SISTER	6.81	2.25	5.37	4.56
287. SKY	7.10	2.17	4.40	3.64
288. SLACK	4.37	4.62	4.88	2.70
289. SLEEP	6.69	2.30	4.80	2.29
290. SLIME	5.29	5.20	4.42	2.94
291. SLOW	4.67	4.21	4.77	1.95
292. SMALL	2.62	3.73	5.27	3.93
293. SMOOTH	4.99	2.83	4.97	2.91
294. SNAIL	6.43	4.29	4.58	1.73
295. SNOW	5.21	2.90	4.98	3.20
296. SOCIALISM	2.78	4.37	3.45	4.10
297. SOFT	5.88	3.15	5.97	2.68
298. SOMBER	3.79	3.81	3.99	2.24
299. SONG	5.63	2.78	4.86	4.28
300. SOUR	3.63	4.77	3.21	4.73
301. SPANKING	4.94	3.99	2.59	5.78
302. SPARKLING	6.72	2.53	4.15	5.85
303. SPICY	5.35	3.32	3.59	5.99
304. SPIDER	4.37	4.47	4.20	5.13
305. SQUARE	4.68	3.68	2.68	3.66
306. STAGNANT	6.63	5.51	4.25	2.13
307. STARS	5.64	2.48	3.68	3.37
308. STARVING	7.37	5.40	3.69	3.71
309. STATUE	6.57	2.70	2.31	2.24
310. STEAL	6.66	5.65	3.82	5.22
311. STEM	4.41	2.92	4.10	3.06
312. STENCH	4.37	4.87	3.03	4.45
313. STIFF	4.77	4.52	2.58	3.50
314. STOP	4.82	3.20	2.69	4.80
315. STOUT	3.49	4.10	4.22	3.11
316. STOVE	5.21	2.77	3.10	4.29
317. STREET	5.59	3.16	2.17	4.71
318. STUDY	6.32	2.41	2.60	4.70
319. SUCCESS	7.67	1.92	2.18	3.50
320. SUDDEN	5.88	4.03	2.81	6.36
321. SUNDAY	5.92	2.38	4.40	2.83
322. SUNLIGHT	7.70	1.86	3.58	4.07
323. SUPERIOR	6.88	2.50	2.20	5.01
324. SWEEPING	4.01	3.10	3.88	5.00
325. SWEET	5.93	2.61	5.12	3.83
326. SWIFT	6.71	2.94	2.38	6.33
327. SWORD	6.91	4.11	1.68	6.03
328. SYMPHONY	6.74	2.13	3.85	4.98
329. TABLE	5.72	3.05	2.37	2.54
330. TALL	5.14	3.05	2.28	4.34
331. TASTE	5.22	2.64	3.84	4.63
332. TENSE	4.58	4.36	3.02	5.37
333 THIEF	6.68	5.59	3.29	5.46
334. THINK	6.29	2.22	3.37	4.59
335. TIME	5.37	2.68	3.21	4.40
336. TORNADO	8.23	4.91	1.80	6.72
337. TOWN	4.28	3.10	3.47	4.70
338. TREES	7.86	2.08	2.15	2.85
339. TROUBLE	5.31	5.03	3.10	5.40
340. TRUNK	4.88	3.34	2.29	2.81
341. TRUTH	7.55	1.91	2.84	4.48
342. UGLY	5.59	5.32	3.50	3.80
343. UNDER	2.38	4.40	3.87	3.55
344. U. N.	5.91	2.34	2.98	4.63
345. UP	4.60	2.80	3.12	4.70
346. VISION	6.32	2.20	3.50	4.87
347. WAGON	4.06	3.70	2.59	3.55
348. WAR	7.85	5.12	1.79	6.21
349. WASH	5.48	2.57	4.19	4.50
350. WATER	4.89	2.77	3.63	4.70
351. WEB	3.89	4.35	4.37	3.24
352. WET	2.64	4.13	4.60	4.24
353. WHITE	4.99	2.74	4.73	3.14
354. WINDOW	5.47	2.82	3.14	3.34
355. WINTER	3.49	3.58	3.10	4.08
356. WISE	6.78	2.01	3.17	4.21
357. WISH	3.83	3.11	4.52	3.86
358. WOMAN	6.43	2.53	5.65	4.55
359. YOUNGSTER	5.42	3.20	4.42	5.90
360. ZENITH	4.20	3.11	2.75	4.88

Factor scores have been computed for the three principal factors described by Osgood, Suci, and Tannenbaum.[5] These factors have been termed 'Evaluation' (Factor I), 'Potency' (Factor II), and 'Activity' (Factor III). It should be noted that the data in the Atlas have not been factor-analyzed and that these scores provide only a rough estimate of the position of the concepts in the semantic space. The scores have been assigned by averaging the values given to the scales having high loadings on the appropriate factors. The scales employed and their loadings on the factors under consideration are given in Table I. The factor scores have been found useful in selecting concepts for further study and experimentation. Concepts with specified characteristics may be located approximately with the factor scores and then scrutinized in detail by reference to the actual profiles and the table of distances.

The data concerning degree of polarity and factor scores are presented in Table II. The first column contains all the concepts from the Atlas in

[5] Osgood, Suci, and Tannenbaum, *op. cit.*, 50-63.

alphabetical order. The second column gives the degree of polarization of each concept. The third column reports the average value for each concept on the eight scales representing Factor I. (Low values are 'good'; high values are 'bad.') The fourth column gives the average value for each concept on the three scales representing Factor II. (Low values are 'potent' and high values are 'impotent.') The fifth column gives the average value for each concept on the three scales representing Factor III. (Low values are 'passive' and high values are 'active.'

RETINAL DISPARITY AND DIPLOPIA VS. LUMINANCE AND SIZE OF TARGET

HAROLD SIEGEL and CARL P. DUNCAN, Northwestern University

In studies of stereoscopic thresholds, the measure taken is typically the smallest amount of retinal disparity that can be detected.[1] For this type of threshold, extremely good acuity-values have been found; most values reported range from 1.5″ to 100″ of arc. Apparently no study has as yet been made to measure the amount that can be tolerated up to the point where diplopia occurs. It is this measure, that is the concern of the present study. As will be seen later, the method used maximized the amount of disparity tolerated. The effect on this measure of two variables, photopic luminance and size of target is reported here.

METHOD

Apparatus. The stimulus-lights were projected on a translucent screen, 28-in. square, by a set of dual projectors (Bell and Howell), each equipped with a tungsten-filament lamp of 300 w. These lamps served as the only light-source.

Four pairs of neutral-density filters, which were held in spectacle frames worn by *O*, were used to vary luminance. By means of an RCA type 930 photoelectric tube, the light transmitted through the filters was measured at the position at which *O* was seated before the screen. The four values of luminance obtained were 0.04, 0.13, 0.38, and 1.13 millilamberts (ml.). All of these values are within the photopic range.

The other independent variable, size of target, was manipulated by means of four sets of black outline-circles. Measured to the outer edge of their contours, the circles subtended visual angles at *O*'s eye of 52′, 1°58′, 5°2′, and 7°3′, respectively. The width of the contour of each circle subtended 18′. Another circle, 11°32′ in visual angle, was used for demonstration.

* Received for publication August 25, 1958.

[1] R. N. Berry, Quantitative relations among vernier, real depth, and stereoscopic depth acuities, *J. exp. Psychol.*, 38, 1948, 708-721; R. N. Berry, L. A. Riggs, and C. P. Duncan, Relation of vernier and depth discriminations to field brightness, *ibid.*, 40 1950, 349-354; C. H. Graham, Visual perception, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 890-891; C. H. Graham, L. A. Riggs, C. G. Mueller, and R. L. Solomon, Precision of stereoscopic settings as influenced by distance of target from a fiducial line, *J. Psychol.*, 27, 1949, 203-207; A. Guggenbuhl, Das stereoskopische Sehen des hell- und dunk eladaptierten Auges, *Ophthalmologica*, 115, 1948, 193-218; H. J. Howard, A test for the judgment of distance, *Amer. J. Ophthal.*, 2, 1919, 656-675; C. G. Mueller, and V. V. Lloyd, Stereoscopic acuity for various levels of illumination, *Proc. nat. Acad. Sci.*, 34, 1948, 223-227.

The two fields, each with its circle, were polarized at 90° and 0° by filters placed in the optical system of each projector. A pair of Polaroid filters crossed so as to separate the vision of *O*'s eyes was mounted in the spectacle frames he wore, together with a pair of spherical convex lenses and prisms to compensate for the fixation-distance used in the experiment. The optical system controlled for accommodation and convergence by placing the apparent image of the patterns at optical infinity for all conditions of viewing.

The experiment was conducted in a dark room. Movement of *O*'s head was minimized by an adjustable head- and chin-rest (Bausch and Lomb) which was mounted on a table 28-in. high. *O* sat in a chair with his forearms on the table as a further aid to stability. *E* adjusted the headrest so *O*'s line of regard was perpendicular to the screen when he looked at a black fixation-cross, ½ in. in width and in height, placed in the center of the outline-circle. The screen was mounted on the table 22.75 in. from *O*'s eyes. The projectors were 29.38 in. behind the screen.

Subjects. The *O*s were 120 men and women, undergraduates from the introductory course in psychology. Replacements were obtained for 7 *O*s who, because of various visual defects, were unable to fuse the targets. As each *O* appeared for the experiment, he was assigned in turn to one of the four conditions of luminance. Thus, to avoid changes in adaptation during the course of the experiment, an independent group of 30 *O*s was tested under each condition of luminance. All the *O*s judged all four sizes of target.

Procedure. All the *O*s were given 16 trials; 4 trials with each of the 4 sizes of circles. Since 4 items (circles) can be arranged in 24 possible orders, 6 sets of 4 orders, each order representing one presentation of each of the 4 circles, were prepared. Each *O* within a block of 6 *O*s went through a different set of 4 orders. Insofar as possible, the four orders assigned to an *O* were such that the same circle did not appear on consecutive trials. The six sets of orders were repeated with each succeeding block of 6 *S*s.

To maximize the amount of retinal disparity for diplopia, every trial was conducted as follows. The trial began with the circles fused. *E* then gradually and continuously separated the two images by turning a dial on the projectors until *O* reported seeing two circles. The images were separated at a rate of about ½° of visual angle per sec. Thus, every *O* made a judgment of diplopia on every trial before a measure (degree of separation of the images) was recorded.

After *O* was seated and the headrest adjusted, he was instructed that upon signal, he was to fixate the cross on the screen. Fixation of the cross without moving head or eyes was emphasized. *O* was told that when he first saw a single circle he was to report this to *E*, and then when he later saw two circles instead of one, he was to report this immediately. The instructions and procedure were clarified for *O* by means of three trials with the demonstration-circle. The 16 experimental trials began immediately after completion of the demonstration, but *O* was given a 2-min. rest between each block of four experimental trials.

As soon as *O* reported seeing double on each trial, *E* recorded the separation of the centers of the circles to the nearest 1/16 in.

Results and Discussion

The mean of the four trials on a particular circle was used as the raw measure of disparity for each *O*. Inspection of the data revealed some positive skew in each of the 16 distributions (4 luminances and 4 sizes of target); therefore the raw scores were converted to logarithms of the distance, which largely eliminated the skewness. Since the set of distributions represent both independent (luminance conditions) and correlated (target-size conditions) data, it was not possible to test for heterogeneity of variance among all 16 distributions simultaneously. The variances of log scores of the four luminance-groups were, however, compared by Hartley's $F_{max.}$ test four times, once for each of the four sizes of circles. The largest of these values was 2.39, which with 4 and 29 *df.* is not significant.

An analysis of variance which permits comparison of both the independent and correlated distributions was performed on the log scores. The analysis is summarized in Table I. It is clear that luminance was not a significant variable (tested against 'between-*O*s-in-same-group' as error); the *F* is less than one. Size, and the interaction of luminance with size, are tested against 'pooled-*O*s-times-size' as error. Table I shows that size of target was highly significant, but the interaction term was not significant. The magnitude of disparity for diplopia increased directly as size of circle increased, but was not significantly affected by luminance within the range used.

TABLE I
Analysis of the Log-Scores for Both Variables

Source	*df.*	*MS*	*F*	*P*
Luminance (L)	3	.031	—	—
Between *O*s in same group	116	.075	—	—
Size (S)	3	2.266	323.71	.001
L×S	9	.003	—	—
Pooled *O*s×S	348	.007	—	—

The mean score for each condition was converted to units of visual angle and is shown as a function of size of target in Fig. 1, with luminance as the parameter. The conclusion from the analysis of variance is equally clear in Fig. 1; the disparity at which diplopia first appears increased directly with circle-size, but was relatively little affected by changes in luminance.

The range of luminance tested here (0.04 to 1.13 ml.) was rather narrow, and it was entirely in the photopic region. Its effect did not reach a significant level. The slight effect that did occur was such that amount of disparity increased as luminance decreased (see Fig. 1). Both of these findings correspond to the results of studies of stereoscopic threshold wherein

luminance was varied.[2] Those studies showed that the threshold decreased (*i.e.* disparity increased) as luminance decreased to rather low values, but the change in threshold was found to be very small within the range of luminance employed.

Target-size had a highly significant effect on the measure of disparity. The mean disparity increased from 1°49′ with the smallest target (52′ circle) to 3°49′ with the largest target (7°3′ circle). This finding corresponds to the results of experiments on stereoscopic threshold where the variable studied was angular separation of the targets.[3] As separation of targets increased, *i.e.* as the stimuli were projected to more peripheral areas

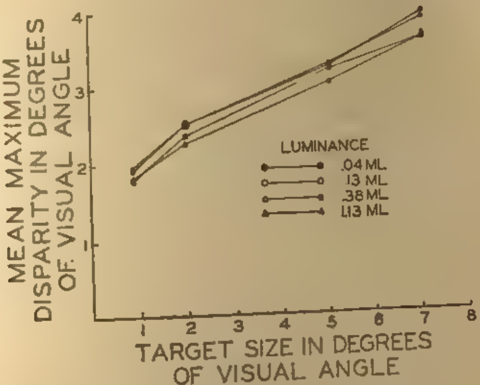


Fig. 1. Mean Retinal Disparity for Diplopia as a Function of Luminance and Size of Target

of the retina, stereoscopic thresholds decreased (again indicated by an increase in values of disparity). The fact that larger values of disparity were found for more peripheral regions of the retina, in both this study and in the studies of stereoscopic threshold, is probably accounted for by the greater anatomical convergence of receptor pathways in more peripheral areas of the retina.[4]

Although the findings of the present study are in the same direction as those studies of stereoscopic threshold that employed the same variables, the order of magnitude of disparity-values found in the present study (2°–4°) is far larger than typical values reported for stereoscopic thresholds

[2] Berry, Riggs, and Duncan, *op. cit.*, 349-354; Mueller and Lloyd, *op. cit.*, 223-227.

[3] Berry, *op. cit.*, 708-721; Graham, *op. cit.*, 890; Graham, Riggs, Mueller, and Solomon, *op. cit.*, 207.

[4] S. L. Polyak, *The Retina*, 1941, 191-449.

(2″–100″). Some of the reasons for this difference are obvious. A deliberate attempt was made to maximize disparity by beginning each trial with the circles fused, then gradually and continuously separating the images until *O* gave a judgment of diplopia. Thus, the images were not separated by discrete steps, with *O* making a judgment at each step, and the measure of disparity was taken only after *O* made a judgment of diplopia on every trial.[5] If, by chance, errors of habituation operated more strongly than errors of expectation, the disparity-values are inflated to that extent. *O*'s reaction-time in reporting diplopia, and *E*'s reaction-time in stopping the separating of the images, undoubtedly operated to increase the disparity that was recorded. Finally, the type of target used may be important; outline-circles may well yield larger values of disparity than stimuli consisting of points or lines.

Summary

The experiment was concerned with the amount of retinal disparity that could be tolerated before fusion was lost and diplopia occurred. On each trial, *O* fixated the center of, and reported initial fusion of, superimposed outline-circles, after which the stimuli were gradually and continuously separated until *O* reported diplopia. Both size of target, and level of photopic luminance, were varied.

The results were that retinal disparity to diplopia did not vary as a function of photopic luminance, but increased directly and significantly with the size of the target. The disparity measured under all conditions was fairly large.

---

[5] In this connection it is worth noting that attempting to proceed in the opposite direction, from diplopia to fusion, would probably increase the variability of the data since, unlike the experience of diplopia, some time is usually required to achieve fusion even when it can be achieved at all. In any case, proceeding from diplopia to fusion would not have yielded the measure that was desired here.

THE EFFECT OF INTERPOLATED, EMOTIONALLY TONED STIMULI ON LEARNING AND RECALL

By Frank A. Scott, East Carolina College

As a review of the litertaure reveals, the interpretation of the effects of interpolated, emotionally toned stimuli on learning and recall is ambiguous. Various experimental results have been reported. Some investigators have found that the recall of the materials originally learned is facilitated;[1] others that it is inhibited.[2]

The purpose of the present study was to devise a series of experiments that would permit an analysis of the data to harmonize the apparently conflicting results previously reported, and to suggest an explanation of the effects of an emotionally toned stimulus on retroactive inhibition.

Various hypotheses regarding this problem may be adopted. The interpolation of emotionally toned stimuli may: (1) increase the tonicity of *S* and thereby facilitate the recall of the material originally learned; (2) interfere with the recall of the original material; or (3) interfere with the learning of the interpolated material and thereby facilitate the recall of the original material.

The third hypothesis, that the emotionally toned stimulus (electric shock) has a detrimental effect on the learning of the interpolated material or upon any activity that immediately follows it, is the one adopted in this study and the one that was investigated.

According to this hypothesis, a shock coming just before the interpolated learning will interfere with the accomplishment of that task and thus will reduce its inhibitory effect upon the recall of the material originally learned. If the shock is introduced between interpolated learning and the

* Received for publication December 24, 1958. This paper is based upon a Master's thesis at Duke University, directed by Dr. Howard Easley.

[1] M. M. White, The influence of interpolated electric shock on recall, *J. exp. Psychol.*, 15, 1932, 752-757; M. E. Bunch and F. D. McTeer, The influence of punishment during learning on retroactive inhibition, *ibid.*, 15, 1932, 473-495; S. H. Britt, Retroactive inhibition: Review of the literature, *Psychol. Bull.*, 32, 1935, 381-440.

[2] L. M. Harden, The effects of emotional reactions upon retention, *J. gen. Psychol.*, 3, 1930, 197-221; J. A. McGeoch, The influence of four different interpolated activities upon retention, *J. exp. Psychol.*, 14, 1931, 400-413; W. D. Tait, The effects of psycho-physical attitudes on memory, *J. abnorm. soc. Psychol.*, 8, 1913, 10-37.

recall of the material originally learned, it should interfere with the recall of the original material.

Procedure. Six different conditions were tested in this study; one was a control and five were experimental. The orginal problem (Problem A) was learning the position of eight chessmen on a chessboard. The interpolated problem (Problem B) was the learning of a different arrangement of the same eight chessmen on the board. The chessmen consisted of king, queen, two castles, bishop, knight, and two pawns. The men were glued to the boards and the same arrangements were used during the entire experiment. As Skaggs showed, retroactive inhibition does occur under these conditions.[8]

Electric shock was the major variable manipulated. As the emotional reaction to shock varies greatly from individual to individual, the experimenter (*E*) believed that better control could be exercised over this variable, not by giving all the *Ss* the same amount of shock, but by giving them as much as they could possibly stand.

TABLE I

DESIGN OF THE VARIOUS EXPERIMENTAL CONDITIONS

(All periods 45 sec. in length.)

Groups	Pre-period	Experimental periods 1	2	3	4	5
Control		learn A	rest	learn B	rest	test A
Experimental 1		learn A	rest	learn B	rest	test B
Experimental 2	shock	learn A	rest	learn B	rest	test A
Experimental 3		learn A	shock	learn B	rest	test A
Experimental 4		learn A	rest	learn B	shock	test A
Experimental 5		learn A	shock	learn B	rest	test B

Shock was interpolated at various positions in the experiment (See Table I). The apparatus consisted of an inductorium and two electrodes wired to two 1½-v. batteries connected in series. The shock was controlled by moving the secondary coil of the inductorium upon the primary. Every *S* controlled the amount of shock he received by releasing one or both of the electrodes. The current was gradually increased until the intensity *S* was willing to bear had been reached, when he terminated the shock. Table I lists the six different conditions studied in this experiment.

The Control Group learned Problem A, rested for 45 sec., then learned Problem B. After another 45 sec. of rest they were tested on Problem A. Experimental Group 1 followed the same procedure as the Control Group but were tested on Problem B. Neither of these first two groups received electric shock. Experimental Group 2 received shock at the beginning of the experimental period and then proceeded with the same schedule as the Control Group.. Experimental Group 3 was confronted with the same problem as the Control Group except that shock instead of a rest-period was interpolated between learning Problem A and Problem B. In Experimental Group 4, shock instead of the rest-period was interpolated between learning

[8] E. B. Skaggs, Further studies in retroactive inhibition, *Psychol. Monogr.*, 34, 1925, (No. 161), 1-60.

Problem B and the test on Problem A. Otherwise, Group 4 was similar to the Control Group. Experimental Group 5 followed the same procedure as Group 3 except that a test was given on Problem B instead of Problem A.

The shocks required approximately 45 sec. to administer. The *Ss* were allowed 45 sec. to learn the position of the chess pieces. There were 45 sec. of undirected rest-intervals between the successive activities when shock was not interpolated at those points. *S* sat in his chair during the course of the experiment.

Instructions. The following instructions were given the *Ss* before the beginning of the experiment:

You will have two problems to learn. You will be given 45 sec. for each of them. Try to learn them both equally well. In learning the problem, I wish you to note the distinction in the shapes of the chessmen and try to remember precisely where the men of the various shapes are. Sometime during the experiment, but not during the actual learning, I will give you an electric shock. You will have perfect control of how much shock you want to take. Do not be afraid for you will not get a sudden shock. The current will be slowly increased until you have had as much as you want. For the experiment to be valid, I want you to take as much shock as you possibly can. It is not dangerous, and when you have had enough, just turn loose of either one or both of the plates. Try to take as much electricity as you can.

TABLE II

Mean Error Scores for Every Group

	Control group	Experimental groups 1	2	3	4	5
Mean error-scores	18.2	16.0	21.1	17.2	18.7	18.8

There were 18 *Ss* for every group, or 108 for the entire experiment. The *Ss*, graduate and undergraduate students at Duke University, were inexperienced with the game of chess.

Results. The results were computed by counting the number of squares that the chessmen were placed out of position at the test. The squares were counted vertically and horizontally, but not diagonally. Table II gives the average scores for each condition. It should be noted that all scores are error-scores and that a lower score represents greater recall and, presumably, less inhibition.

Although the differences among the six groups were small and not significant, the consistency of the results throughout the experiment with our hypothesis makes them more significant than their individual reliabilities warrant. That the electric shock is detrimental to whatever follows it can be seen by comparing the various scores. The worst score was made by Group 2, which had shock at the beginning of the experiment. This group made a mean score of 21.1 as compared with the Control Group whose

score was 18.2. Shock seems to have had also a detrimental effect upon recall. In Group 4, which received shock immediately preceding the recall of Problem A, a mean score of 18.7 was made as compared with the Control Group score of 18.2. In Group 3, which received shock immediately preceding the learning of Problem B, a better score was made on the recall of Problem A than was made by the Control Group: 17.2 as compared with 18.2. In the groups which were asked to recall Problem B (Groups 1 and 5), the one that had no shock made the better score; 16 as compared with the group with the shock which made a score of 18.8.

Discussion. The conditions of the Control and Experimental Group 3 are similar to those used by White, and our results are similar to those obtained by him.[4] The better scores obtained by Group 3 do not, however, seem to be due to increased tonicity of the *Ss*, as White suggested, but to a decrease in the level of the learning of the interpolated material. This decrease is shown by comparing the scores of Group 5 with those of Group 1, 18.8 vs. 16. The results of this study conflict with those of Harden and McGeoch.[5] Those experimenters, however, spaced the shock at 30-sec. intervals throughout the interpolated periods which meant that within 30 sec. before the *Ss* were tested on the original learning, they received a shock. This shock was deterimental to recall and, consequently, increased the amount of measured retroactive inhibition. Such results are found when the scores of the Control Group are compared with the scores of Group 4, 18.2 vs. 18.7.

Minami and Dallenbach, using cockroaches as subjects, studied the effects of forced activity on learning and retention.[6] Their results were interpreted in terms of two factors—an anti-consolidation factor which was apparently retroactive in its effect, and a factor of excitement or irritability. Both factors were detrimental to learning and retention as measured by relearning. The present study is not designed to show the anti-consolidation factor as reported by them. In comparing Group 3 with the Control Group (mean of 17.2 as against 18.2), there is no indication that the shock in Group 3 interfered with the consolidation of the original learning. The anti-consolidation factor could be present, however, in all six conditions of this experiment, as the longest time-interval between activities was 45 sec.

The factor of excitement can explain the differences between the groups of this study. This factor has proactive effect on anything that follows its arousal. This proactive effect is illustrated when the performances of Groups 2 and 5 are compared with the performance of the Control Group and Group 1, respectively.

Bunch and McTeer used electric shock as punishment for errors in both the original and interpolated material.[7] They found that the punished individuals learned faster and made fewer errors than the unpunished group. Their experiment probably influenced Britt to conclude that elcetric shock in either the original or the inter-

[4] White, *op. cit.*, 752-757.
[5] L. M. Harden, The effects of emotional reactions upon retention, *J. gen. Psychol.*, 3, 1930, 197-221; McGeoch, 14, 1931, 400-413.
[6] Hiroshi Minami and K. M. Dallenbach, The effect of activity upon learning and retention in the cockroach, this JOURNAL, 59, 1946, 1-58.
[7] Bunch and McTeer, *op. cit.*, 473-495.

polated learning tended to decrease retroactive inhibition.[8] Bunch and McTeer's experiment is, however, hardly comparable to the others dealing with the effect of electric shock on retroactive inhibition as these experiments used shock during the actual learning. This introduced another variable, the varying similarity of the original and interpolated learning problems.

Reynolds discovered that a second interpolation of learning destroyed the inhibitory effects of the first interpolation, thus permitting better recall of the original learning.[9] His results are consistent with the general theory that whatever interferes with an inhibitor thereby reduces its inhibitory effect. This theory is illustrated when Group 3 scores are compared with Control Groups scores (mean of 17.2 vs. a mean 18.2). That shock did interfere with the learning of Problem B is seen in the comparison of scores of Group 1 (16.0) with those of Group 5 (18.8).

Conclusions. The results of this experiment tentatively indicate that the introduction of an emotional stimulus may either increase or decrease recall-scores. These effects are entirely consistent with the results of experiments previously reported. There is no evidence, however, of an emotional stimulus' having any direct facilitating effect on either learning or recall. The effect of the stimulus seems to be primarily (if not entirely) proactive rather than retroactive. It is detrimental to both learning and recall, when these activities immediately follow it.

[8] S. H. Britt, *op. cit., Psychol. Bull.*, 32, 1935, 381-440.

[9] Bradley Reynolds, The mnemonic function of interference, *J. exp. Psychol.*, 30, 1942, 336-341.

A FURTHER TEST OF 'REASONING' IN RATS

By John P. Seward, Robert B. Jones, and Stanley Summers, University of California, Los Angeles

In Maier's three-table test, or its T-maze equivalent, a rat is first allowed to explore an empty maze with distinctive end-points, then placed directly at one such point with food. The question is whether it later can choose the path leading to the food-place.[1] As an experiment of 'latent learning,' this design seemed once to provide a crucial test of cognitive as opposed to S–R theories of learning. It is now clear that neither doctrine is likely to stand or fall by the outcome. Yet there is still the problem: what are the necessary conditions of 'reasoning' as here defined? Interest is sharpened by the results to date: of a dozen published experiments, seven are positive and five negative.[2]

Recent theorizing by MacCorquodale and Meehl and by Spence has revived the inquiry in a slightly different form.[3] It was the purpose of the present experiment to try out a relatively new technique for demonstrating the acquisition and use of expectancies by rats.

Briefly, the method has three parts. In Part 1, a hungry rat is put in an enclosed runway and makes one or more trips to an empty end-box. In Part 2, it is put directly in the end-box and given a pellet of food. In Part 3, it is again placed at the entrance of the runway. By the assumptions of expectancy-theory, the rat should run faster on the first trial of Part 3 than at the end of Part 1.

To tighten the argument, several sources of error must be avoided. First, the end-box must be quite different from the runway to minimize generalization from one to the other. Secondly, the effect of *eating in the end-box* must be separated from that of eating in any box and from that of just being in the end-box without food. Control of these factors suggests a 2 × 2 design.

Method and Procedure

Apparatus. The runway was 9½ ft. long, 4 in. wide, and 3 in. high, separated from a start-box, 9½ × 4 × 3 in., by two vertically sliding doors, one opaque, the other transparent. Runway and start-box were of wood, painted black, and roofed with hardware cloth. Two photocells, 6 in. from the start-door and 12 in. from the end, controlled a 1-r.p. sec. and a 6-r.p.m. timer, respectively. The end-box was of unpainted pine, 10 × 6¾ × 10¾ in., with a roof of window-screen. It stood be-

* Received for publication February 26, 1959. This work was supported by a grant from the Research Committee of the University of California.

[1] N. R. F. Maier, The effect of cerebral destruction on reasoning and learning in rats, *J. comp. Neurol.*, 54, 1932, 45-75.

[2] See Kenneth MacCorquodale and P. E. Meehl in W. K. Estes *et al.* (eds.) *Modern Learning Theory*, 1954, 211; E. R. Hilgard, *Theories of Learning*, 2nd ed., 1956, 213; J. P. Seward, Basic issues in learning theory, in G. H. Seward and J. P. Seward (eds.), *Current Psychological Issues: Studies in Honor of Robert S. Woodworth*, 1958, 128-132.

[3] MacCorquodale and Meehl, *op. cit.*, 237-247; K. W. Spence, *Behavior Theory and Conditioning*, 1956, 124-164.

side the runway, from which it was entered through a door cut in one side at the far end, and from which it could not be seen. The plywood door was black on the side facing the runway and unpainted on the other. All experimenting was done in a darkened room. Light was provided by five 7.5-w. lamps mounted 16 in. above the apparatus; curtains of cheesecloth on both sides formed a 'one-way screen' and shut out most extra-maze visual cues.

A 'neutral' wooden box, 16 × 4 × 5½ in., stood about 5 ft. to one side of the runway. It was painted gray and had a hardware-cloth roof and a transparent door at one end. Overhead lighting and cheesecloth drapes gave considerable one-way screening. Both the end-box and the neutral box contained a glass furniture-coaster.

Subjects. The *Ss* were 71 Sprague-Dawley female rats 3—4 mo. old at the outset. Sixty-four survived matching and completed the experiment in two replications of 36 and 28 *Ss*.

Maintenance. The animals were housed in cages of four or five with water always available, but without food. Once a day, at least 15 min. after experimenting, they were fed Steenbock mash for 40 min. Adaptation began after about one week on hunger-cycle.

Procedure. (*a*) *Adaptation: Day 1.* All *Ss*, in groups of two or three, were given 6 min. to explore the end-box (with door closed) and the neutral box. *Day 2.* All *Ss* alone explored end-box and neutral box for 3 min. *Day 3.* All *Ss*, 4 or 5 at a time, explored the apparatus with all doors open for 12 min. They also had 3 min. alone in the neutral box. *Days 4-5.* A random half of the *Ss* had 6 min. to explore the apparatus singly, while the other half had 3 min. in the neutral box. Next day the treatments were reversed.

(*b*) *Training: Part* 1. All *Ss* had one timed trial a day for 12 days. They were run in a random order changed daily. Procedure in a single trial was as follows: *E* put *S* in the start-box. When *S* faced the opaque door, *E* raised it; at the count of three he raised the transparent door, automatically starting both timers. *S* was removed 30 sec. after entering the end-box, the door of which was left open.

Part 2. The *Ss* were ranked for median total times to the second photocell and grouped in levels according to speed; four levels in the first replication, five in the second. Within levels the *Ss* were assigned randomly in equal numbers to four groups for differential treatment. Counting both replications, there were 16 *Ss* in each group.

Treatments started the day after the last timed trial. They consisted of one trial a day for three days as follows: Group *EF*. After baiting the coaster with a 1-gm. pellet of mash, *E* put *S* directly in the end-box with door closed and removed it after it had eaten the pellet. Group *NF*. In the same way each *S* received one pellet in the neutral box. Group *EO*. After removing all traces of food, *E* put *S* directly in the end-box for 1 min., the approximate median eating time for the *F*-groups. Group *NO*. Similarly, each *S* spent 1 min. in the neutral box without food.

(*c*) *Test: Part 3.* Starting the following day, all *Ss* were given one timed trial a day for six days. Conditions were the same as in Part 1 except that the end-box was baited with a 1-gm. pellet.

Scoring. For the purpose of normalizing the distributions, all times were converted into reciprocals. Two scores were obtained for each *S* on each trial after the eighth day of training: (*a*) *starting speed*—10 times the reciprocal of the time in seconds from opening the door to stopping the first clock; and (*b*) *running speed*—100 times the reciprocal of the time in seconds between the two photocells.

Results

Table I shows mean running speeds on the last four training trials of Part 1 and the first test-trial of Part 3. The experimental group, *EF*, ran more rapidly on the test than before; all three control groups ran more slowly.

To evaluate the significance of these changes in speed we first did an analysis of variance of the data represented by Table I. The interaction between rows and columns was significant at the 2.5% level. For the difference within a single group to reach the 5% level (two-tailed), a difference of 5.0 was required. By this criterion Group *NF*'s loss was significant but Group *EF*'s gain was not. The sign test, however, pointed in the opposite direction: of the 16 *Ss* in Group *EF*, 13 increased their speed ($p = 0.022$); of the 15 in Group *NF* that changed, 10 were slower ($p = 0.30$). By still a third criterion, Wilcoxon's signed-ranks test, both differences were significant at the 5% level.

TABLE I

Mean Running Speeds on Trials 9–12 of Training and Trial 1 of Test

Trials	Group			
	EF	*NF*	*EO*	*NO*
Training 9–12	21.2	21.5	20.6	20.8
Test 1	26.0	15.5	16.4	17.8
Diff.	4.8	−6.0	−4.2	−3.0

Comparing the data within rows of the table, we find the group equally matched at the end of training. On the first test-trial, analysis of variance showed them to be significantly different ($p = 0.01$); the factorial interaction, too, was significant ($p = 0.025$). By *t*-tests, Group *EF* proved superior to *NF* and *EO* at the 1% level and to *NO* at the 2% level (two-tailed); other differences were negligible. (On Test-Trial 2, it should be noted, the four group means came together and remained so for the rest of the experiment, reaching an apparent asymptote on Day 5.) Reciprocals of starting time were analyzed in the same way as those of running time, but no group-differences were found.

Discussion The results for Test-Trial 1 suggest the following interpretation: Our *Ss* presumably acquired an expectation of the end-box in Part 1. In Part 2, the experimental group, but not the controls, formed an association between end-box and food. When the *Ss* of this group were returned to the runway in Part 3, their anticipation of the end-box mediated a new expectancy with motivating properties; this would explain their faster running. The absence of effect on starting-speed might be due either to a weak expectancy in the start-box owing to the remoteness of the goal, or to lack of time for the expectancy to function.

Spence's theory of incentive motivation is quite similar to the one just presented.[4] He, too, holds that an ante-dating goal-reaction, r_g, aroused in the start-box, energizes the locomotor response. He differs from us, however, in making the occurrence of r_g in the start-box contingent on generalization from the goal-box. He there-

[4] Spence, *op. cit.*, 124-164.

fore assumes a goal-box "highly similar in its physical properties to the runway."[8] For us this requirement is unnecessary, since we assume that a reponse to the goal-box itself can be learned, carried forward, and used to mediate *rg*.

It is unlikely that Spence's theory as it stands would fit these data. One would have to assert that the end-box, but not the neutral box, generalized to the runway. True, the design would have been improved if the two boxes had changed places for half the *Ss*; but in size and color the neutral box was more similar than the end-box to the runway, and stimulus-generalization should therefore have favored Group *NF* rather than *EF*. Extra-maze cues could have operated only while the *Ss* were being transferred to and from the apparatus; otherwise they were largely eliminated by one-way screening. Even if we assume that room-cues received in transit gave Group *EF* an advantage, it is hard to see how they could have done so by direct stimulus-generalization, since the room-cues at the end-box were no more similar to the cues at the start than were those at the neutral box.

Two experiments recently have been performed to test Spence's theory under conditions designed to permit generalization, and both gave essentially negative results.[9] This is surprising in view of our data, since if 'latent learning' can be demonstrated *without* generalization it certainly should appear when that factor is added.

Here is a possible solution to the paradox: As the authors of both papers suggested, generalization from goal-box to runway actually may have interfered with the predicted result by causing *S* to seek food instead of run. There should be less interference from an *rg* aroused by anticipatory reactions to a differentiated goal-box. This interpretation may also explain a secondary finding of the present experiment. Apparently, eating in the neutral box retarded Group *NF* about as much as eating in the end-box accelerated Group *EF*. If, as we have contended, there was more generalization to the runway from neutral box than from end-box, Group *NF* may have been distracted by a stronger tendency to seek food in the runway.

The only way to settle these questions, of course, is by an experiment expressly designed to separate the effects of generalization from those of 'expectancy' on the measured response.

Summary. A modified form of 'reasoning' test was made in three parts. In Part 1, 64 rats were given one trial a day for 12 days in a runway to a dissimilar, empty end-box. They then were matched for speed and assigned to four groups. In Part 2, Group *EF* was put directly in the end-box and given a pellet of food once a day for three days. Group *NF* was similarly fed in a neutral box. Groups *EO* and *NO* were put in end-box and neutral box, respectively, without food. Part 3 consisted of one trial a day for six days to a pellet in the end-box.

On the first day of testing, Group *EF* ran significantly faster than the other three groups. Compared with their previous performance, Group *EF* ran faster and Group *NF* slower; these changes were of border-line significance. No other dependable differences were found. The main finding was attributed to incentive motivation mediated by expectancy of the end-box.

[8] Spence, *op. cit.*, 147.

[9] Larry Stein, The classical conditioning of the consummatory response as a determinant of instrumental performance, *J. comp. physiol. Psychol.*, 50, 1957, 269-278; C. F. Swift and E. L. Wike, A test of Spence's theory of incentive motivation, *Psychol. Rec.*, 8, 1958, 21-25.

FIGURAL AFTER-EFFECTS WITH A STABILIZED RETINAL IMAGE

By JOHN KRAUSKOPF, Brown University

Hochberg and Hay recently tested the proposition that movements of the retinal image of the inspection-figure (*I*-figure) resulting from physiological nystagmus and other eye-movements are necessary for the establishment of figural after-effects.[1] By presenting the *I*-figure very briefly, and allowing *O* to observe its after-image, they obtained a kind of 'stopped image.' Under these conditions subsequently viewed test-figures (*T*-figures) appeared altered in size in the same way as they would under conditions generally used for the study of figural after-effects. The authors concluded that "physiological nystagmus is probably not *necessary* for figural after-effects; we cannot yet say whether it is sufficient."[2]

These conclusions are well put. In the theory proposed by Osgood and Heyer to account for the figural after-effects, physiological nystagmus plays a dual role: (1) It results in a spread of the effective retinal image; and (2) it provides the necessary conditions for the activation of 'on-off' and 'off' elements in the regions of the retina on which the contours of the figure play.[3] The second presumed effect of physiological nystagmus is more important theoretically for it provides a mechanism whereby the locus of a contour would be registered in the same way when direction of the brightness-gradient at the contour is reversed. The second role of physiological nystagmus is crucial because experimental results reported by Köhler and Wallach indicate that the direction and magnitude of figural after-effects are independent of the direction of the brightness gradient.[4]

Hochberg and Hay found it necessary to have the screen on which *O* projected the after-image intermittently illuminated that the after-images would remain visible throughout the inspection-period (*I*-period). These

* Received for publication May 27, 1959. The author is now at Rutgers University.

[1] J. E. Hochberg and John Hay, Figural after-effect, after-image, and physiological nystagmus, this JOURNAL, 69, 1956, 480-482.

[2] Hochberg and Hay, *op. cit.*, 482.

[3] C. E. Osgood and A. W. Heyer, A new interpretation of figural after-effects, *Psychol. Rev.*, 59, 1951, 98-118.

[4] Wolfgang Köhler and Hans Wallach, Figural after-effects, *Proc. Amer. Phil. Soc.*, 88, 1944, 269-357.

conditions would in all likelihood result in the activation of 'on-off' and 'off' elements. Since the retina is differentially adapted by the brief stimulus-flash, the strength of the 'on-off' and 'off' activity would not be uniform over the retina. Although the pattern of activity is different from that obtained with the orthodox procedure, it does seem that the Osgood-Heyer model might predict the results obtained. The present experiments were therefore designed to test the effects of motion of the retinal images on the production of figural after-effects in such a way that both effects of eye movements were controlled.

METHOD

In the present experiments, measurements of the magnitude of figural after-effects were made under normal conditions of fixation and, as in the Hochberg and Hay experiments, under conditions which prevented motion of the retinal images. Motion of the retinal image was eliminated by the technique known as the 'stabilized image.'[5] The essential feature of this technique is that the image viewed by *O* is made to move in direct correspondence with eye-movements and thus remains fixed in place on his retina. This is accomplished by reflecting the light from the stimulus-projector by means of a mirror embedded in a contact lens worn in the observing eye and forming an image with this light on a viewing screen. By appropriate adjustment of the optical distances from the mirror to the screen, and from the screen to the viewing eye, the movement of the image on the retina may be stopped completely. In previous experiments, only horizontal eye-movements were completely compensated; in the present experiment, both vertical and horizontal stabilizations were obtained.[6] The most striking effect of stabilizing an image is that, although it appears initially to be as sharp and clear as a similar image normally viewed, it soon fades and disappears. If, however, as in the present experiments, the target is relatively large and the contrast high, it will remain visible for a considerable period. Although the targets used in the present experiment did disappear briefly during the inspection-period, no special techniques were employed to keep them visible.

Stimuli. The stimuli used are illustrated in Fig. 1. The *T*-figures are given in Fig. 1 A, and the *I*-figure in Fig. 1 B. Fig. 1 C is a composite of the two, and indicates the relations between them. The figures appeared on a white circular background 11.5° in diameter which had a brightness of 1800 ft.-L. The black rectangles (2° × 1°) were produced by interposing opaque templates in the light-beam. The inner edges of the *T*-rectangles and the *I*-rectangle were 2° from the fixation-point. The difference in height of the lower edge of the *I*-figure and the upper edge of the

---

[5] L. A. Riggs, Floyd Ratliff, J. C. Cornsweet, and T. N. Cornsweet, The disappearance of steadily fixated visual test objects, *J. opt. Soc. Amer.*, 43, 1953, 495-501; R. W. Ditchburn and B. L. Ginsborg, Vision with a stabilized retinal image, *Nature* (Lond.), 170, 1952, 36-37.

[6] R. W. Ditchburn and D. H. Fender, The stabilized image, *Optica. Acta*, 2, 1955, 128-133.

left T-rectangle was 12′. The right T-rectangle could be moved vertically by a micrometer-control. The small, bright fixation point was seen under normal viewing conditions. The presentation of the T- and I-objects was controlled by a synchronous-motor timer. In Experiment I, these objects were seen under stabilized-image conditions; in Experiment II under normal conditions of fixation.

Procedure. Except for viewing conditions, the procedure was the same in both experiments. An experimental session was divided into three parts. In the first part, a pre-inspection determination of the point of subjective equality (*PSE*) for the height of the rectangles was made by the method of limits. The T-figure was presented once every 15 sec. for 0.5 sec.; the rest of the time the field was dark except

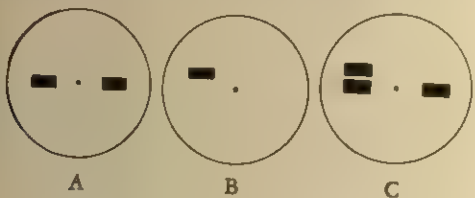


FIG. 1. THE STIMULUS-OBJECTS
(A = T-figure; B = I-figure; C = composite.)

for the fixation-point. In the second part of the experiment, a high level of 'satiation' was produced by presenting the I-figure for 3 min. In the third part of the experiment, the post-satiation *PSE* was determined. The 15-sec. cycle was used again, but this time the I-figure was presented for 13.5 sec. followed by a blank interval of 0.5 sec., then the test-flash, again 0.5 sec. in duration. A cycle ended with another blank interval of 0.5 sec. The purpose of presenting the I-figure during the last part of the experiment was to maintain a high level of satiation. This procedure has been found to give large and stable after-effects.[7] Two measurements of pre- and post-satiation *PSE*s were made in each session (one ascending and one descending series). Each experiment consisted of eight such sessions. The sessions were separated by at least 1 hr.

Since the stabilized-image technique requires considerable training and individual fitting with contact lenses, only one *O* (the author) was used.

RESULTS

Figural after-effects were obtained in both experiments, *i.e.* the right rectangle was set in a lower position for the post-inspection *PSE* than for the pre-inspection *PSE*. The direction of the difference between the pre- and post-satiation *PSE*s was the same for all eight sessions of both experiments and thus reliable ($p < 0.01$) according to a sign test. The mean

[7] John Krauskopf, The magnitude of figural after-effects as a function of the duration of the test-period, this JOURNAL, 67, 1954, 684-690.

after-effect for a stabilized image (Experiment I) was 5.6′ while for normal fixation (Experiment II) it was 5.3′. The *SD*s of the difference in *PSE*s were 4.1′ and 3.6′ for stabilized viewing and normal fixation, respectively. The difference between the effects in Experiments I and II was not reliable ($t = 0.14$). The results seem to indicate that the elimination of retinal-image motion does not influence figural after-effects.

Conclusions

In the two experiments the stimulus-conditions differed only in respect to the movement of the image across the retina. According to the model proposed by Osgood and Heyer, it is the 'on-off' and 'off' activity generated at contours as a result of retinal-image motion which is responsible for figural after-effects. The results show that eliminating the conditions for this activity does not eliminate the after-effect. The results indicate therefore, that the Osgood-Heyer theory is invalid.

THE DISTANCE-PARADOX IN KINESTHETIC FIGURAL AFTER-EFFECTS

By JOHN KRAUSKOPF and TRYGG ENGEN, Brown University

Since the appearance of Köhler and Wallach's monograph on figural after-effects in two dimensional vision,[1] similar effects have been demonstrated in the visual third dimension, in kinesthesis, and in audition.[2] Köhler and Wallach demonstrated that a simple adaptation-model, such as that of Gibson,[3] is not sufficient to account for the evidence on figural after-effects. The so-called distance-paradox is one of the most important differences between adaptation-effects and figural after-effects. Through adaptation, change in apparent magnitude is a monotonic function of the magnitude of the adapting stimulus, *i.e.* the more intense the adapting stimulus, the greater the effect. Köhler and Wallach found that visual figural after-effects increase at first, eventually reach a maximum, and then diminish as the distance between the inspection- and test-objects increases.

The present paper is concerned with the distance-paradox in kinesthetic after-effects. The distance-paradox, which has been shown in other modalities, is theoretically important because it seems to depend on place rather than intensitive neural representation.[4] Moreover, it is well defined quantitatively and applicable to any perceptual dimension. In addition to its possible relevance for neurophysiology, evidence concerning the distance-paradox would be interesting from a purely formal, psychophysical point of view.

Method. The *Os* were 60 college students of both sexes with no previous experience with figural after-effects. The experiment was a regular part of the laboratory of an elementary psychology course.

The task of *O* was to find with the pads of his thumb and middle finger that width of a variable wooden bar which appeared equal to a 2-in. standard bar before and after the presentation of an inspection-bar. The variable bar varied in width from 1 1/16 to 3 1/16 in., in 1/8-in. steps (1/16 in. on each side and 17 steps in all). The bars were 17 in. long and were mounted 2/3 in. above the table-top. Before entering the experimental room, *O* was blindfolded. The two bars were placed before *O*

* Received for publication March 23, 1959.

[1] Wolfgang Köhler and Hans Wallach, Figural after-effects, *Proc. Amer. phil. Soc.*, 88, 1944, 269-357.

[2] Wolfgang Köhler and D. A. Emery, Figural after-effects in the third dimension of visual space, this JOURNAL, 60, 1947, 159-201; Wolfgang Köhler and Dorothy Dinnerstein, Figural after-effects in kinesthesis, in *Miscellanea Psychologica Albert Michotte*, 1947, 196-220; J. A. Deutsch, A preliminary report on a new auditory after-effect, *Quart. J. Exp. Psychol.*, 3, 1951, 43-46; F. N. Jones and E. B. Bunting, Displacement after-effect in auditory localization, *Amer. Psychologist*, 4, 1949, 389; John Krauskopf, Figural after-effect in auditory space, this JOURNAL, 67, 1954, 278-287.

[3] J. J. Gibson, Adaptation, after-effect and contrast in the perception of curved lines, *J. exp. Psychol.*, 16, 1933, 1-31.

[4] Köhler and Wallach, *op. cit.*, 297 ff.; Jones and Bressler, *op. cit.*, 389; Krauskopf, *op. cit.*, 286.

about .24 in. apart. Judgments were obtained by the method of limits, and five ascending trials were alternated with five descending trials. Half of the *Os* in each group used their left hands, and the other half their right hands on the standard bar. The pre-inspection judgments were followed immediately by the inspection-period of 1 min. during which *O* moved the thumb and middle finger which had previously been used with the standard bar along an inspection-bar. The *Os* were divided into five groups at random. The inspection-bars used for the five groups were 2.5, 3.0, 3.5, 4.0, and 4.5 in. wide. Post-inspection judgments were then obtained for the 2.0-in. standard with the same procedure used in the pre-inspection period.

Results. The measure of figural after-effect was the difference in point of subjec-

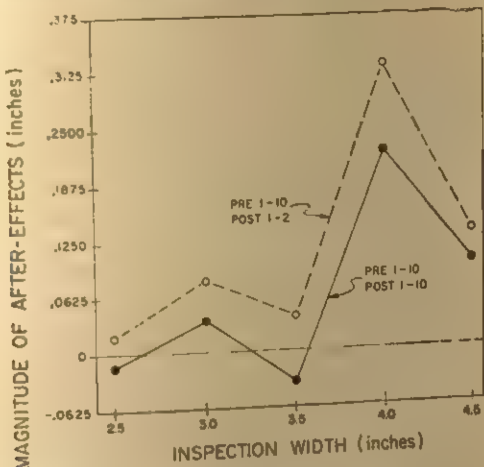


FIG. 1. MAGNITUDE OF KINESTHETIC AFTER-EFFECTS AS A FUNCTION OF THE WIDTH OF THE INSPECTION-BAR

tive equality (*PSE*) obtained for the pre- and post-inspection judgments. Since all the inspection-bars were wider than the standard, a decrease in *PSE* for the post-inspection judgments was positive evidence. The mean *PSE* for the five ascending and five descending trials were computed first to test for anticipatory errors. The mean differences were 0.108 in. and 0.109 in. for the pre- and post-inspection judgments, respectively, and *t*-tests indicate that both are reliable ($p < 0.01$). The pre-inspection judgments were obtained under uniform conditions for all *Os*, while the post-inspection judgments were not, but in both cases the ascending means were lower than the descending means. This anticipatory error was expected on the basis of the nature of the after-effect and was balanced out for the remaining analyses by treating the judgments in pairs of ascending and descending trials.

Two measures of change in *PSE* were computed, the mean difference between (1) all 10 pre-inspection judgments and all 10 post-inspection judgments, and (2) all 10 pre-inspection judgments and the first 2 post-inspection judgments. The results of these analyses are presented in Fig. 1, which shows that the distance-paradox

was obtained because the magnitude of the after-effect tends to increase with the wider inspection-bars up to 4.0 in., but then diminishes again with the 4.5-in. inspection-bar. Analyses of variance were performed on both measures and the F-ratio in each case indicated that the effect of width of the inspection-bar was reliable ($p < 0.01$). The standard errors of the means and the mean differences in *PSEs* were computed from the error-variances. The after-effects obtained for each width of the inspection-bar were tested and those for the 4.0- and 4.5-in. bars were found to be reliable ($p < 0.01$). The differences in the magnitude of the after-effects for neighboring widths of the bars were tested, and those between the bars 3.5 and 4.0 in. and between the 4.0- and 4.5-in. bars were found to be reliable ($p < 0.01$, except for the comparison between the 4.0- and 4.5-in. bars when all trials were included, when $p < 0.05$).

The differences in the magnitude of the after-effects found with the two indices

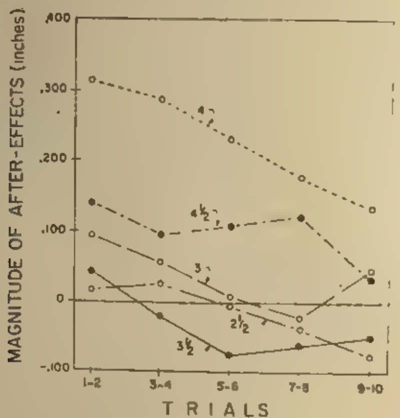


FIG. 2. DECAY OF KINESTHETIC AFTER-EFFECTS

suggest a decay-process. Evidence of decay is presented in Fig. 2, where the means of the 5 pairs of ascending and descending trials for the post-inspection judgments are compared with the means of the 10 pre-inspection judgments. An analysis of variance of the difference in means for the first and last two judgments shows that the effect of width of the inspection-bar on the magnitude of the decay is not reliable ($p > 0.05$), but the standard error of the means estimated from the error-variance indicates that the over-all average decay is reliable ($p < 0.05$). Separate analyses for each inspection-bar indicate reliable decays for the 3.5- and 4.5-in. bars ($p < 0.05$) and for the 4.0-in. bar ($p < 0.01$). These results are consistent with those presented above, and show why larger after-effects are obtained for the first two post-inspection trials. Eliminating the decay over the last eight trials produces a more orderly graph.

Discussion. The present experiment corroborates the earlier evidence of kinesthetic after-effects obtained by Köhler and Dinnerstein and, in addition, reveals further

similarities between the after-effects in kinesthesis and those found in other perceptual dimensions.[5] The decay of the effects seems to be quite rapid; no attempt was made to investigate long-term effects. Our findings are in accord with measurements of the same sort in two-dimensional visual perception and auditory localization.[6]

Köhler and Dinnerstein pointed out that small after-effects should result when *O* moves his hand along the variable bar during the experiment in the pre- or post-inspection periods.[7] They did not present data on this point, but it is clear from the nature and direction of the after-effect obtained that ascending series, *i.e.* when *O* starts with a small width on the variable bar, should give lower *PSE*s than descending series. This prediction was tested and confirmed with the present data. Since an adaptation-model would also predict this anticipatory error, the finding has no apparent theoretical significance and is only of importance for the treatment of data.

The main finding of the present experiment is that the distance-paradox discovered in vision and audition also appears in kinesthesis. The importance of this phenomenon would seem to lie not only in its formal similarity to effects obtained in other perceptual continua but in the possible implications for the neural mechanisms underlying these dimensions. There has been much controversy about the relative merits of different neurophysiological models proposed to account for the visual after-effects.[8] Regardless of specific details, however, these models have been based on the assumption that the effects are due to some form of lateral interaction in a topological map of the retina somewhere in the visual projection system. It has been suggested that the existence of similar effects in other modalities might be taken as evidence that some form of place-localization exists in the processes related to these perceptual continua.[9] Place-localization has been demonstrated for auditory pitch, and Jeffress has suggested a mechanism whereby the difference in the time of arrival of stimuli at the two ears may be represented by place in a localization-'map.'[10] Much less seems to be known about the cortical representation of kinesthetic stimulation, but recent neurophysiological data suggest that limb-position may be represented cortically by place.[11] The present evidence of the distance-paradox may be taken as presumptive evidence for spatial representation of kinesthesis.

[5] Köhler and Dinnerstein, *op. cit.*, 201 ff.

[6] E. R. Hammer, Temporal factors in figural after-effects, this JOURNAL, 62, 1949, 337-354; Krauskopf, *op. cit.*, 286.

[7] Köhler and Dinnerstein, *op. cit.*, 201.

[8] Köhler and Wallach, *op. cit.*, 315 ff.; C. E. Osgood and A. W. Heyer, Jr., A new interpretation of figural after-effects, *Psychol. Rev.*, 59, 1952, 98-118.

[9] Krauskopf, *op. cit.*, 297.

[10] L. A. Jeffress, A place theory of auditory localization, *J. compar. physiol. Psychol.*, 41, 1948, 35-39.

[11] V. B. Mountcastle, Modality and topographic properties of single neurons of cat's somatic sensory cortex, *J. Neurophysiol.*, 20, 1957, 408-434.

APPARATUS

SILVER-SILVER CHLORIDE SPONGE ELECTRODES FOR SKIN POTENTIAL RECORDING

By D. N. O'Connell and Bernard Tursky, Massachusetts Mental Health Center

During research over the last two years involving psychophysiological measurements, it has been desirable to make measurements of absolute level of skin-potential as well as changes in level in response to stimulus-conditions. It was found that the accuracy of potential recording was hampered by the inadequacy of the types of electrodes than currently in use. Such electrodes showed drifts and large biases.

In an attempt to overcome these difficulties, the properties of the non-polarizable silver-silver halide electrode widely used as a reference electrode in electrochemical measurements were investigated. This type of electrode seemed to have the basic characteristics required. A modification of this type of electrode had already been used by Köhler *et al.*[1] for recording cortical potentials.

An electrode rugged enough for both research and routine clinical use that could be prepared by a relatively simple, standard, and economical procedure was desired. Thus, we wished to avoid a liquid contact medium and a mercury terminal, which they used. The following is a description of a silver-silver chloride electrode that has given very satisfactory results in over two years of polygraphic recording.

The electrode has three parts (see Fig. 1): (1) an outer casing of durable plastic; (2) the inner core of silver-silver chloride; and (3) the metal terminal assembly to which the recording leads may be attached. The component parts will be described in order.

(1) Casing. The casing is machined from 0.75 in. clear-cast lucite rod. The lower half of the casing is 1 cm. high and contains a central chamber 7 mm. in diameter and 8 mm. deep, which holds the contact medium. The upper half of the casing is

* This research was supported in part by a Public Health Service Research Grant No. N2820 (A) from the Institute of Mental Health, Public Health Service, and in part by a grant from the Society for the Investigation of Human Ecology.

[1] Wolfgang Köhler, Richard Held, and D. N. O'Connell, An investigation of cortical current, *Proc. Am. philos. Soc.*, *96*, 1952, 290-330.

turned down to equal in diameter the head of the terminal stud (6.6 mm.). This provides a column of convenient diameter for use with standard *EKG* straps. The center column is drilled and threaded to a depth of 8 mm. to fit the terminal stud, and a center hole through which the platinum wire of the core may be led is drilled with a No. 60 drill.

(2) Inner core. The silver-silver chloride core is prepared by a modification of methods described by Janz and Taniguchi.[2] They are of the thermal-electrolytic type. Lengths of platinum wire, B. & S. Gage No. 25, are cut 3 cm. long, cleaned in HCl, dried, and sealed into previously cleaned 6 cm. lengths of soft glass tubing that 2 cm. extends outside the tubing and the portion inside the tubing can make contact with mercury poured into the tubing. The end of the wire outside is looped at its end and again cleaned. This is then dipped in a paste of purified silver oxide along the entire length of the protruding wire and reduced in an electric oven at 400°C. Care must be taken to coat the entire length of the wire. This procedure of

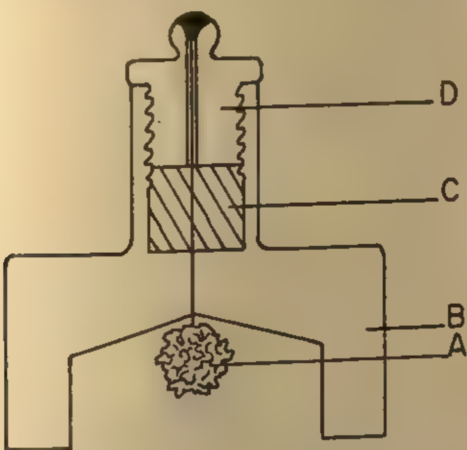


FIG. 1. CROSS-SECTION OF THE ELECTRODE

A = inner core of chloridized silver sponge with central platinum wire; B = clear lucite outer casing; C = adhesive cement sealing core from terminal attachment above it; D = terminal stud attachment.

coating the platinum wire with silver has an advantage over coating it by electrodeposition from a potassium silver cyanide solution in that contamination by cyanide ion, requiring long washing, is avoided. The loop at the end of the wire is then dipped in the silver oxide paste and reduced as before. This step may be repeated until a silver sponge is built up of the desired size, about 5 mm. in diameter. The electrodes are then chloridized in a 0.75 *M* solution of hydrochloric acid for four hours at a current of 4 m.a. per electrode, using a platinum wire as cathode. They

[2] G. J. Janz, and Harry Tanaguchi. The silver-silver halide electrodes: preparation, stability, and standard potentials in aqueous and non-aqueous media, *Chem. Rev.*, 53, 1953, 397-437.

are then washed with distilled water and dried. The glass tubing is then carefully broken away from the platinum wire, after which they are ready to be sealed into the electrode casings.

(3) Terminal attachment. The terminal attachment is made with a 7/16 in. Nu-Way Snap stud, which has been center-drilled with a No. 60 drill and cut down in length to about 3 mm. The platinum wire of the silver-silver chloride core is threaded up through the center hole in the casing. Before screwing on the terminal stud, the center hole is sealed with a drop of Sauereisen Insa-Lute adhesive cement, which is allowed to harden for 24 hrs. The terminal stud is then screwed tightly into the center column, taking care to thread the platinum wire of the core through the hole previously drilled in the stud. A small amount of adhesive cement placed on the under side of the top of the terminal stud before it is screwed down effects a permanent seal and strengthens the terminal assembly. The platinum wire, which extends somewhat above the terminal stud is trimmed down and soldered. The central chamber of the casing is now filled with electrode paste, or whatever contact medium is being used and stored in a small plastic box, which has been provided with holes fitted with ⅜-in. outside-diameter grommets through which the center column of the electrode fits snugly. This method of storage prevents the electrode paste from drying out and facilitates handling of the electrodes. Attachment of recording leads to the studs may be made by using Nu-Way Snap Terminals. Electrode pairs must always be stored shunted. We find that they come into equilibrium in a day or two and may then be paired off.

Twenty-four electrodes were prepared for testing. Tests were made in saline electrode paste and recorded on an Offner Type R Dynagraph, at an amplification of 0.5 mv./cm. The electrodes were randomly divided into 12 pairs. In all cases, no drift was apparent over a 1-hr. recording period, which would indicate any drift present to be below ±0.01 mv./hr.

Electrode bias was determined for all possible combinations of pairs ($N = 276$). The range of these measures, disregarding sign, was from 0.00 to 0.73 mv. The mean bias, again disregarding sign, was 0.17 mv. These bias potentials would be generally considered negligible at the levels of amplification used in skin-potential recording. We have found the electrodes to be stable within this range for periods of a year or more if properly stored.

It is thus apparent that this type of electrode has the characteristics desired for accurate recording of skin potentials: low bias potential, freedom from drift, and long-term stability.

APPARATUS NOTES

A NEW TACHISTOSCOPE FOR ANIMALS AND MAN

The tachistoscope described here is suitable for use with lower species of animals, with primates, and with human *Ss*. Its simplicity, accuracy, flexibility, and nominal cost warrant its report.

The tachistoscope has three essential parts: (1) lamps, daylight and monochromatic, to illuminate the exposure-field; (2) specially prepared stimulus-cards; and (3) a timing device. The basis of the apparatus consists in selecting inks for the preparation of the stimuli that provide these properties: (1) Under ordinary white light, the stimuli are discriminable on the basis of hue- and brightness-differences; while (2) under nearly monochromatic illumination these differences are absent. By introducing white light briefly, while the stimulus-cards are illuminated by monochromatic light, tachistoscopic exposure is provided.

Preparation of the stimulus-cards. Two-dimensional patterns are the most easily prepared stimuli. They may be prepared of any inks which have equal reflectance of a monochromatic source, but they also should afford good contrast for relatively easy discrimination with ordinary light (see Fig. 1). One combination of inks that

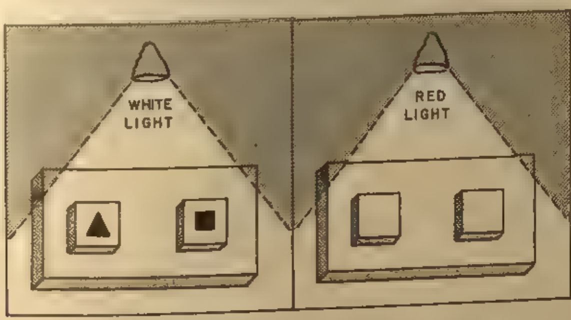


FIG. 1. APPEARANCE OF STIMULUS-CARDS IN RED AND WHITE ILLUMINATION

satisfies these requirements (when the monochromatic light is deep red) is a yellow ink, General Printing Ink No. U65567, and a red, Levey No. 12N230.[1] These are regular commercial inks readily available through a local printing firm. Using these inks, stimulus-cards may be prepared in quantity by lithographic methods, or individually by ordinary drawing techniques.

Lamps. The most convenient source of deep red, nearly monochromatic, light is a red, darkroom safelight. One or more 25-w., ruby red lamps will sufficiently illuminate the exposure-field. The brief exposures of white light may be provided

[1] The designation of suitable inks was provided by Dr. S. M. Newhall, Color Technology Division, Eastman Kodak Company, Rochester, N.Y.

by several types of daylight lamps. An accurate fluorescent lamp system, which can be built for less than $5 (exclusive of the timer), is diagrammed in Fig. 2.[2] This system, by providing continuous current to the heaters, allows instant flashing of ordinary fluorescent lamps. In combination with a suitable timing device, exposures as brief as 0.02 sec. are feasible.

Applications. Preliminary testing has confirmed the suitability of this apparatus

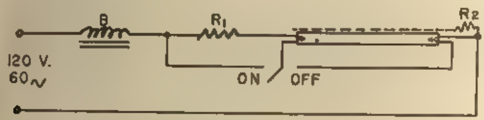


Fig. 2. Circuit for Flashing Fluorescent Lamps

For 14, 15, or 20-w. fluorescent use for 'B' GE ballast 89G381 and for 'R_1' a 200Ω 20-w. resistor. To secure reliable starting, a metal strip at least ⅛ in. wide (broken line) touching the lamp and running its length must be used. The metal strip should be grounded or connected to one of the line conductors through a 1 meg. resistor ('R_2'). It is necessary to energize the circuit with the switch in the 'off' position for at least 2 sec. prior to starting. The circuit is not recommended for general purposes because of possibly reduced lamp-life which is not important in this application.

with monkeys. Its application to other species is, in principle, virtually unaffected by the physiological properties of the animal's visual system.

In use with human *Ss*, it has these advantages over the Dodge tachistoscope: (1) there is no confining case; (2) an entire room can be used as a tachistoscope, allowing group experiments; and (3) relative location of fixation-point and discriminanda may be at virtually any distance and degree of parafoveality.

VA Hospital, Durham, North Carolina

Isaac Behar

[2] This circuit, developed at the General Electric Company, was provided by Dr. S. K. Guth, Radiant Energy Effects Laboratory.

NOTES AND DISCUSSIONS

ON THORNDIKE'S "CONFIRMING REACTION"

In 1933 Thorndike felt compelled to make a statement about the nature of reinforcement and its place in his own scheme. The original Law of Effect, with some modifications and in the face of some "objections to the hypothesis that a satisfying after-effect of a mental connection works back upon it to strengthen it" (an objection which now seems empirically indefensible), has successfully defended its place over a long period.[1] At all events, Thorndike apparently felt justified in writing some 16 yr. later that "the Law of Effect, that the immediate consequence of a mental connection (in particular, a satisfying state of affairs following a connection and belonging to it) can work back upon it to strengthen it, is now accepted by most psychologists."[2] In this era of information-theory and servomechanism, the possibility of the control of behavior according to the information communicated by its consequences can scarcely be questioned. Instances are numerous of error-compensating behaviors, of self-regulations in terms of feedback into the behavior-system of bits of information regarding results up-to-date all tending to stabilize the system-responses, testify to the regulation of behavior by its consequences.

Control by results is no great problem. The difficulty lies with the "satisfying state of affairs" (or "annoying state") which is the immediate consequence of responding. To avoid any ideational connotation of the information supplied by consequences, Thorndike rested his case on objective behavior. He stated:

> By a satisfying state of affairs or satisfier is meant one which the animal does nothing to avoid, often doing things which maintain or renew it. By an annoying state of affairs is meant one which the animal does nothing to preserve, often doing things which put an end to it."[3]

This statement merely restates the empirical Law of Effect—there are some conditions that gain acceptance while others are rejected, which states the objective facts of the matter. The argument goes on:

> A satisfier *exerts an influence* that strengthens any modifiable C, *i.e.* any activity, state, or condition of N, the neurones, upon which this influence impinges."[4]

Thorndike has at this point gone beyond the observables to locate the

[1] E. L. Thorndike, *Selected Writings from a Connectionist's Psychology*, 1949, 13.
[2] *Idem.*
[3] *Ibid.*, 14.
[4] *Idem.*

mechanism of satisfiers and annoyers in the nervous system. He refers to an "unknown reaction of neurones" as an integral part of the reinforcement-process, a reaction which strengthens reactions upon which it impinges. This neural reaction he calls the "yes-reaction," the "OK reaction," or the "confirming reaction."

Obviously presenting a vague, speculative notion, Thorndike was, as always, tentative in his physiologizing and left, as he admits, "its intimate histological basis and physiological nature" out of his account as "no better known than those of facilitation, inhibition, fatigue."[5] Notwithstanding the difficulty of specifying the precise physiology of the reaction, Thorndike refers to some features of the confirming reaction which could be regarded as hypotheses to be confirmed or rejected as neurological evidence arose to support or weaken them. At the time of their formulation these hypotheses could scarcely be tied together by Thorndike's "unknown reaction of the neurones." Nevertheless, the features as he conceived them were stated as follows:

(1) The confirming reaction is independent of sensory pleasure.
(2) The confirming reaction seems often to issue from some overhead control in N, the neural basis of some want or 'drive' or purpose or then active self of the animal.
(3) The potency of a confirming reaction may bear little relation to the intensity of the satisfier. . . . There seems to be an upper point beyond which increases in a reward add only excitement. Toward the low end there is a range where the reward fails more and more frequently to arouse an adequate confirming reaction.
(4) Satisfaction sets the force in action and the force acts on the connection which was just active in intimate functional association with the production of the satisfier, or its near neighbors.[6]

The confirming reaction came under fire as vague, unspecified, and highly speculative. Experiments directed to the study of the behavioral implications of the concept, particularly those of the 'spread-phenomenon,' have remained inconclusive. Yet it is interesting, in the quest for some direct support of the Thorndike hypothesis, to reexamine his conceptualization in the light of recent neurophysiological studies of behavior, especially those executed by Olds and others. These would seem to reinstate the Thorndike view of reinforcement, or at least to give it a further chance.

Using an implanted micro-electrode with rats placed in a Skinner box, in which the lever operated a switch to produce a mild electric shock, Olds and Milner[7] showed that certain nuclei in the septal area, the posterior hypothalamus and other nuclei about the midline of the cerebral hemispheres, produced, when stimulated, observable reinforcing effects char-

[5] Thorndike, *op. cit.*, 16.
[6] *Idem*, 16-19.
[7] James Olds, Pleasure centers in the brain, *Scientific American*, 195, 1956, 107-108.

acteristic of the usual 'rewarding' and 'punishment' situations.[7] Other nuclei did not mediate these effects. Typical instrumental conditioning-effects were observed. If the rat, placed in the box, pressed the lever which delivered a mild shock to an appropriate 'reward' nucleus, it continued to press the lever, often at an amazingly high frequency. The usual associated conditioning phenomena, *e.g.* extinction-effects, obtained. That is, direct electro-stimulation of certain neural centers produced rewarding or reinforcing effects, stamping in a criterion-response. Likewise the stimulation of other areas produced an inhibiting effect upon lever-pressing. 'Satisfiers' and 'annoyers,' as objectively defined by Thorndike, therefore, have demonstrable neural correlates. In other words, there are brain centers whose direct activation by an instrumental response has the effect of satisfiers or annoyers upon that response. Thorndike contended:

> If a S–R connection [or, alternatively, the R–S connection which should be added] has a satisfying after-effect which causes some control in the *N* to send forth a confirming reaction, and if the *S* continues, the confirming reaction tends to cause a continuance or continued repetition of the *R* then and there, and often with more vigor and shorter latency.[8]

Several studies testify to the vigor of the animals' responses; Olds reported response-rates up to 7000 per hr.[9] Evidence for latency-effects is not so clear, since the studies reported do not appear to have been directed to this response-variable. A study which involved time spent in a runway suggests, however, that electro-stimulated rats are at no disadvantages as compared with food-rewarded rats.[10] There is evidence, therefore, of 'reward' and 'punishment' centers in the brain. In support of Thorndike's views about the unequal effects of these two motivators, especially of his later statements ascribing the major reinforcing effects to the positive reward, Olds may be quoted:

> The areas in which the stimulation produces the approach or rewarding effect occupy a larger proportion of the brain than do the areas in which the avoidance or punishing effect is produced.

[As an argument against the *drive-reduction*, in contrast to the theory of reinforcement (confirming reaction), he continues:]

> Therefore, the brain cannot be thought of as tending mainly to produce behaviors which *decrease* its own excitation, for a large portion causes behaviors which *increase* excitation.[11]

---

[8] Thorndike, *op. cit.*, 17.
[9] James Olds, Self stimulation of the brain, *Science*, 127, 1958, 315-324.
[10] Olds, Runway and maze behavior controlled by baso-medial forebrain stimulation, *J. comp. physiol. Psychol.*, 49, 1956, 507-512.
[11] Olds, *op. cit.*, *Science*, 127, 1958, 318.

Turning now to the four behavioral features of the confirming reaction proposed by Thorndike, we find that a compelling verification of his speculations reveals itself in the literature on intra-cranial self-stimulation.

(1) Since the self-stimulation involves the direct activation of prosencephalic nuclei, it is difficult to see how the exteroceptive sensorium could be causally involved. It is impossible to say how the animal 'feels' under conditions of intracranial self-activation. Presumably 'conscious feels' could be mediated by interoceptive and proprioceptive afferents, but again it is not easy to see how these are directly involved in the intra-cranial excitation effect. It is probably safer and certainly more parsimonious to conceive, as Thorndike did, of an activity in one part of the brain exerting an influence which is capable of intensifying activities going on concurrently in another part of the brain. Otherwise, a sensory hedonism involving the mediation of 'ideas' or 'feels' would seem to be perilously close at hand.

(2) The evidence accumulating around the relationship between basic organismic needs and self-stimulation is quite spectacular and supports Thorndike's hypothesis of 'overhead control.' Not only are the reinforcing effects of self-stimulation more effective when combined with a basic need such as food-deprivation, but the most recent explorations support the view that there are centers associated with the specific needs. Thus Olds, in his latest review, writes:

> Further studies indicate that the electrical brain-shock reward has the effect of a primary reward object in several different experimental situations. These studies suggest also that the electric brain shock excites cells which are normally involved in the mediation of the effects of conventional primary reinforcers such as food and sex objects.[12] [In the summary of this review he is even more explicit:] "These primary reward systems of the brain are subdivided into specific drive-reward sub-systems mediating the specific drives such as hunger and sex."[13]

In the interests of brevity references to the specific experiments are not made in the text, but the interested reader may follow up the empirical findings from the literature cited below.[14] It should be noted that the behaviors studied in the intracranial self-stimulation experiments extend beyond the Skinner box to runways and mazes and to *S*s other than rats. The results in all cases support the general effect of reinforcement by brain-stimulation.

(3) Thorndike's third feature refers to thresholds and the effects of varying the intensity of rewards. Deliberate studies involving the employment of increasing

[12] Olds, *op. cit.*, *Science*, 127, 1958, 318.

[13] *Ibid.*, 324.

[14] In addition to the works already cited, the reader is referred to: J. M. R. Delgado, W. W. Roberts, and N. E. Miller, Learning motivated by electrical stimulation of the brain, *Amer. J. Physiol.*, 179, 1954, 587-593; Murray Sidman, J. V. Bardy, J. J. Boren, D. G. Conrad, and A. Schulman, Reward schedules and behavior maintained by intracranial self-stimulation, *Science*, 122, 1955, 830-831; J. M. R. Delgado, H. E. Rosvold, and Edmund Looney, Evoking conditioned fear by electrical stimulation of subcortical structures in the monkey brain, *J. comp. physiol. Psychol.*, 49, 1956, 373-380; J. V. Brady, J. J. Boren, Donald Conrad, and Murray Sidman, The effect of food and water deprivation upon intracranial self-stimulation, *ibid.*, 50, 1957, 134-137; H. P. Zeigler, Electrical stimulation of the brain and the psychophysiology of learning and motivation, *Psychol. Bull.*, 54, 1957, 363-382; Ben Bursten and J. M. R. Delgado, Positive reinforcement induced by intracerebral stimulation in the monkey, *J. comp. physiol. Psychol.*, 51, 1958, 6-10.

current as an independent variable have supported Thorndike's mechanism from the neurophysiological point of view. Allowing for differences between centers of reward and punishment, and for differences within these main classes of areas, it is clear that the micro-stimulation must reach a threshold value to be effective. In connection with increasing rewards, Reynolds reports a study of the behavior of seven rats under conditions of varying voltage with rate of hypothalamic self-stimulation as the dependent variable.[15]

> The most important result was that although there was an increase in rate of self-stimulation as voltage was increased, this occurred only up to a maximal value, beyond which further increases in voltage produced a decline in rate. This was true of all seven rats under all the motivational conditions, and is contrary to Olds's observation that rate will continue to increase with voltage-increases up to and even beyond the seizure-thresholds.

The precise state of affairs here is not determinable from the available reports.

(4) Finally, with respect to the action of "the force on the connection which was just active," a return to the original observations of Olds will supply the empirical neuro-psychological support. In Olds's own words:

> We were not at first concerned to hit very specific points in the brain, and in fact in our early tests the electrodes did not always go to the particular areas in the mid-line system at which they were aimed. Our lack of aim turned out to be a fortunate happening for us. In one animal the electrode missed its target and landed not in the mid-brain reticular system but in a nerve pathway from the rhinencephalon. This led to an unexpected discovery.
>
> In the test-experiment we were using, the animal was placed in a large box with corners labeled A, B, C and D. Whenever the animal went to corner A, its brain was given a mild electric shock by the experimenter. When the test was performed on the animal with the electrode in the rhinencephalic nerve, it kept returning to corner A. . . At this point we assumed that the stimulus must provoke curiosity—we did not yet think of it as a reward. Further experimentation on the same animal soon indicated to our surprise that its response to the stimulus was more than curiosity. On the second day, after the animal had acquired the habit of returning to corner A to be stimulated, we began trying to draw it away to corner B, giving it an electric shock whenever it took a step in that direction. Within a matter of five minutes, the animal was in corner B. After this, the animal could be directed to almost any spot in the box at the will of the experimenter.[16]

In other words, the rat reproduced the response that *had just been rewarded,* or there is a confirming reaction on the connection that has just been active.

Quite apart from the tremendous significance of the intra-cranial self-stimulation studies of Olds and others in their own right, as contributions to the neurophysiological correlates of behavior, there seems to be some point in referring them back to the older speculative theory, if only as a reminder that the speculations of the past often become the empirical data of the present. Over all of this, however, shines the genius of the man who conceived what might be the case, and whose vague speculations seem to

---

[15] R. W. Reynolds, The relationship between stimulation-voltage and rate of hypothalamic self-stimulation in the rat, *J. comp. physiol. Psychol.*, 51, 1958, 193-198.

[16] Olds, *op. cit., Scientific American,* 195, 1956, 107-108.

have been intelligent anticipations. This is a clear example of a conceptualization dictated by the observable behavioral phenomena well in advance of the possibility of direct empirical confirmation. Yet, like Kubie's theoretical reverberatory circuit, later verified by Lorente de No experimentally, Thorndike's investment in the confirming reaction would appear to be paying off handsomely.

The University of New England
New South Wales, Australia
P. E. H. Barratt

PAVLOV, BECHTEREV, AND TWITMYER ON CONDITIONING

Dallenbach has raised several interesting questions in his Note, "Twitmyer and the conditioned reflex,"[1] which referred to my Note, "Salivary conditioning before Pavlov."[2] Pointing out that I had limited my review to salivary conditioning, Dallenbach stated, "By so doing, Twitmyer's study of the knee-jerk, in which the discovery of the conditioned response was announced and the phenomenon first described, escaped his attention."[3] This, Dallenbach suggested, was understandable since Twitmyer's study was not listed in any of the annual indexes of psychology. "A scholar . . . would search in vain for references to Twitmyer's work."[4]

Twitmyer's study was included in an early draft of my Note and was later eliminated because it falls outside the scope of a review on salivary conditioning. The study is referred to by Hilgard and Marquis and also by Woodworth.[5] Thus, while it is true that Twitmyer's work did not attract much attention, it has not entirely escaped attention.

Whether Twitmyer was the first to discover and describe conditioning is a question that deserves consideration. Dallenbach states that Pavlov reported his results in 1904 while Twitmyer reported his, in his doctoral dissertation, in 1902. To call attention to his findings, Twitmyer gave a paper at the meeting of the American Psychological Association in 1904, but his listeners showed no sign of interest. Dallenbach comments, "Had Twitmyer received a spark of encouragement, he would have continued his investigation. Had he done that, 'conditioning' might have had its effective beginning in America instead of Russia."[6] It is well to remember that

[1] K. M. Dallenbach, Twitmyer and the conditioned reflex, this Journal, 72, 1959, 633-638.

[2] M. R. Rosenzweig, Salivary conditioning before Pavlov, this Journal, 72, 1959, 628-633.

[3] Dallenbach, *op. cit.*, 634.

[4] Dallenbach, *op. cit.*, 637.

[5] E. R. Hilgard and D. G. Marquis, *Conditioning and Learning*, 1940, 3; R. S. Woodworth, *Experimental Psychology*, 1938, 114 f.

[6] Dallenbach, *op. cit.*, 636, footnote 8.

Pavlov's work on conditioning grew out of his extensive work on digestive secretion and that he described 'psychic stimulation' of salivary and gastric secretion in his book of 1897.[7] Responses to 'psychic stimulation' had to be brought under control for a thorough study of digestive secretion. Thus Pavlov's interest in conditioning not only antedated Twitmyer's but it was also closely related to his whole program of research. Pavlov was therefore convinced of the importance of the phenomenon and was ready to study it thoroughly. Gantt has described some of Pavlov's early steps:

> He had noticed, and then he began to track down, the irregularity that occurred in the gastric and salivary secretions . . . sometimes this secretion appeared before the meat was given to the dog. By his observation he found out what the source of the discrepancy was; the dog began secreting not only when he saw the man who was accustomed to feeding him but also when he heard his footsteps. That, then, was the origin of his work on the conditional reflex.
>
> These observations were made before the turn of the century, but only after 1900 did he begin a systematic study of the so-called conditional reflex.[8]
>
> Pavlov listed the first reports on conditioning from his laboratory as follows: The first paper giving a more or less systematic description of 'natural' conditioned reflexes—then called 'psychic reflexes'—was published by Dr. Wolfson in the form of a thesis entitled "Observations upon salivary secretion" (Petrograd, 1899). The term 'conditioned reflex' was first used in print by Dr. Tolochinov, who completed his experiments in 1901 and communicated the results at the Congress of Natural Sciences in Helsingfors in 1903.[9]

Actually Tolochinov, a young psychiatrist, delivered his paper to the Helsingfors Congress in 1902 and it appeared in print the following year.[10] Thus by 1902 Pavlov's group had already given their discovery a distinguishing name; such a name, Dallenbach has suggested, is an important means of calling attention to a discovery. Pavlov first gave a paper on conditioning at the International Congress of Medicine, Madrid, in 1903.[11] This paper was published in Russia in the same year and in a German journal the following year.[12] A more detailed account, giving

[7] I have consulted the German translation which appeared the next year: J. P. Pawlow, *Die Arbiet der Verdauungsdrüsen*, trans. by A. Walther, 1898, 84-116.

[8] W. H. Gantt, Pavlov, in M. A. B. Brazier (ed.) *The Central Nervous System and Behavior*, 1959, 169.

[9] I. P. Pavlov, *Conditioned Reflexes*, trans. by G. V. Anrep, 1927, 412.

[10] I. F. Tolotschinoff, Contribution a l'étude de la physiologie et de la psychologie des glandes salivaires, *Nord. natur. läkeremötet*, 8, 1903, 42-46. For more information about Tolochinov see: A. K. Federova-Grot. The controversy between Dr. I. F. Tolochinov and Academician I. P. Pavlov in 1912-1914, *Sechenov Physiol. J. U.S.S.R.*, English trans., 43, 1957, 452-455. After the initial work on conditioning, Tolochinov, a young psychiatrist who had taken his doctorate with Bechterev, turned to full-time clinical practice. Pavlov was unable to persuade him to extend and to publish his studies. Ten years later, without consulting Pavolv, Tolochinov published a series of aritcles on his earlier work giving his own interpretations which differed in important respects from those of Pavlov. Pavlov disavowed these articles but later continued to speak warmly of Tolochinov.

[11] Pavlov, *Lectures on Conditioned Reflexes*, trans. by W. H. Gantt, 1928, Chap. 1.

[12] J. P. Pawlow, Psychische Erregung der Speicheldrüsen, *Ergeb. Physiol.*, 3, 1904, 177-193.

protocols of several experiments, appeared in French in 1904.[13] Several collaborators are mentioned in this paper as investigating various aspects of conditioning. It is clear that the study of conditioning was well under way in Russia in the first years of the century. It seems doubtful that, even if Twitmyer had received encouragement at the meeting of 1904, " 'conditioning' might have had its effective beginning in America instead of in Russia."

Gantt's account accords with Dallenbach's picture of Pavlov as "a man who could not be discouraged by an adverse intellectual climate." Gantt depicts as follows Pavlov's dilemma in the 1890's:

> Over a period of several years, Pavlov considered whether he would make a physiological attack on the problem of irregularities of secretion, or whether he should neglect them. This was a very difficult period for him because during this time he was advised by both Tigerstedt and Sherrington to stop this fad of psychic secretion and return to real physiology. It was against the course of physiology at that time, both in Russia and abroad, to study this so-called psychic secretion.[14]

The conditioning of *somatic* or motor responses, first described by Twitmyer, was brought into prominence by the research and writings of Vladimir M. Bechterev.[15] Bechterev made important contributions in neuroanatomy, neurology, and psychiatry. His later work was devoted to establishing 'reflexology' as the objective science of personality from the bio-social viewpoint. In 1886-87 Bechterev had found that lesions in the motor cortex of the dog could abolish learned movements, such as presenting the front paw when demanded by the trainer. Then, beginning in 1898, there was a series of contributions from his laboratory indicating that "natural, associative reflexes," both motor and secretory, could be abolished by localized cortical lesions. After Pavlov began reporting on salivary conditioning, Bechterev did some experiments along the same line, but he was not completely satisfied with this method. He writes:

> Being dissatisfied with this method, especially because of its inapplicability to man, I made a communication, in the spring of 1907, to the Society of Physicians of the Hospital of Nervous Diseases, and proved, on the basis of experiments made by me in collaboration with Dr. Spirtov, the possibility of producing an artificial association-motor reflex in the respiration of a dog. (See *Minutes* of these Meetings.)
>
> A little later, the same reflex was produced in my laboratory on a man also (Dr. Anfimov), and then an artificial association-motor reflex was produced in my laboratory by the electrical stimulation of a dog's paw (Dr. V. P. Protopopov), and this experiment essentially improved the methodology of investigation of association-

[13] Pawlow, Sur la sécrétion psychique des glandes salivaires, *Arch. internat. Physiol.*, 1, 1904, 119-135.

[14] Gantt, *op. cit.*, 169.

[15] For a recent sketch and appraisal of Bechterev, see P. I. Yakovlev, *Bechterev*, in M. A. B. Brazier (ed.), *The Central Nervous System and Behavior*, 1959, 187-210.

motor reflexes in dogs. After this, the artificial inculcation of an association-motor reflex from the sole, and later from the fingers, of a man . . . was secured in my laboratory (Dr. Molotkov, *Dissertation*, St. Petersburg) according to a method suggested by me.

From this time, progress in the elaboration of this subject began to be swift as a result of the establishment of a new method, and gradually it became possible in my laboratory, to produce in man an artificial association-motor reflex to a sound—an association-motor reflex in the form of the reflex which is obtained by knocking with a little hammer on the patellar ligament (Dr. Shevalev), and another reflex to a sound—a reflex after electric stimulation of the palm or the finger-tips (Bechterev and Shtechelovanov).[16]

Thus, as the last paragraph shows, conditioning of the knee jerk was rediscovered independently in Bechterev's laboratory. There is no indication that Bechterev ever knew of Twitmyer's work.

The translation of Bechterev's book, *Objective Psychology,* into French and German in 1913 helped to bring conditioning to the attention of western scientists.[17] An entire chapter was devoted to 'association reflexes,' both motor and secretory. It was the French translation that provided the guide for the experiments of Watson and Lashley at Johns Hopkins in 1915. Watson devoted his presidential address to the American Psychological Association in December 1915 to an enthusiastic report of the possibilities of this new method.[18] Only eleven years after Twitmyer's paper had failed to gain any response at the psychological meetings, the phenomenon he had discovered was at the center of the stage, but under the names of Pavlov and Bechterev.

Watson shared Bechterev's opinion that motor conditioning provides greater scope and is more applicable to human experimentation than is salivary conditioning. Nevertheless he used the Pavlovian designation "conditioned reflex" rather than Bechterev's term, "association reflex." Hilgard and Marquis have suggested, "It is probably to [Bechterev] more than to Pavlov that we owe the bold acceptance of conditioning by psychologists, although the details of conditioning which came to be accepted were Pavlov's."[19]

Writing these comments gives me an opportunity to amplify the material concerning Robert Whytt in my previous Note. A recent opportunity to consult the original work showed that he had in 1763 anticipated later concepts of conditioning more

---

[16] V. M. Bechterev, *General Principles of Human Reflexology,* trans. by Emma and William Murphy, 1933, 197-198.

[17] W. Bechterew, *La Psychologie Objective,* trans. by N. Kostyleff, 1913, 249-267. W. von Bekhtereff, *Objektive Psychologie; oder, Psychoreflexologie; die Lehre von den Assoziations-Reflexen,* 1913.

[18] J. B. Watson, The place of the conditioned-reflex in psychology, *Psychol. Rev.,* 23, 1916, 89-116.

[19] Hilgard and Marquis, *op. cit.,* 9.

completely than I had realized from a secondary account. Here is one important passage.

> We consider, that not only an irritation of the muscles of animals, or parts nearly connected with them, is followed by convulsive motions; but that the remembrance or *idea* of substances, formerly applied to different parts of the body, produces almost the same effect, as if those substances were really present. Thus the sight, or even the recalled *idea* of grateful food, causes an uncommon flow of spittle into the mouth of a hungry person; and the seeing of a lemon cut produces the same effect in many people. . . . The sight of a medicine that has often provocked [sic] vomiting, nay the very mention of its name, will in many delicate persons raise a nausea.[20]

University of California MARK R. ROSENZWEIG

WEBER'S LAW AND THICKNESS DETERMINED TACTUALLY

The tactile perception of thickness follows Weber's law.[1] The procedure used to demonstrate this is simple and available to every laboratory. *S* is blindfolded and seated in a quiet room facing the investigator. A number of consecutive leaves from a book with smooth edges and even binding is offered to *S* in one hand and he is instructed to select with his other hand the same number of leaves by their thickness in some other part of the book. He is allowed 30 sec. for his judgment, after which the book is removed and another group of consecutive leaves from another book of the same issue and condition is offered. The position of the sample in the book and the number of leaves of the sample are changed randomly to avoid a bias of worn corners or progressive changes in size, although the latter was tested and found not to be a serious problem. No significant differences were noted between continuing descending and ascending sample-sizes. No information as to sample-size or accuracy is given *S* during the experiments.

It will be noted that the determinations of this study do not precisely follow Weber's law, but represent a modification of it. Weber's law would require that the measurements be taken at the point of noticeable difference, *e.g.* at the point where one sample is just larger or just smaller than the other. The *Ss* were instructed to add or to turn back leaves until the second sample was equal the first. Their judgments, except when equal,

---

[20] Robert Whytt, *An Essay on the Vital and Other Involuntary Motions of Animals,* 2nd ed., 1763, 280.

[1] T. H. Howells, Is Weber's law reducible to the physical coefficient of friction?, *J. Gen. Psychol.*, 50, 1954, 249-260; W .C. Michels and Harry Helson, A reformulation of the Fechner Law in terms of adaptations level applied to rating-scale data, this JOURNAL, 62, 1949, 355-368; Jean Piaget, B. von Albertini, and M. Rossi, A probability interpretation of Weber's law and the laws of relative fixations, *Arch. Psychol.*, 30, 1944, 95-138.

were greater or less than the comparison group—but not by a just noticeable difference. My hypothesis is that the *average error* is a constant function of the sample size, rather than the *just noticeable difference* of Weber's law.

The results shown in Fig. 1 are based on 20 determinations of 9 *S*s of 17 standard groups of leaves. The plot for each sample-size gives the standard deviation and standard error for each of the 20 determinations. The line is fitted by the method of least squares.

A difference was noted between preferred and non-preferred hands. Fig. 2 shows the curves for determinations made by the right and left hand

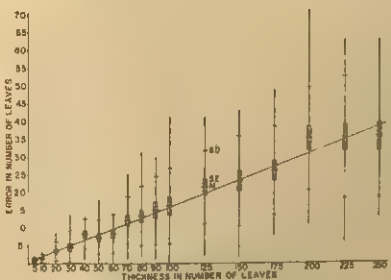


FIG. 1. TACTUAL THICKNESS: ERROR AS A FUNCTION OF SAMPLE SIZE

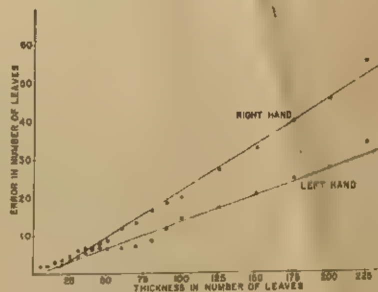


FIG. 2. TACTUAL THICKNESS: ERROR AS A FUNCTION OF SAMPLE SIZE (*S*s is Left-Handed)

for a left-handed *S*. It is interesting to see that both curves indicate a constant ratio, even though the left hand is more accurate.

An attempt was made to investigate the degree of error the *S*s would make when told to make the variable set 'just smaller' and 'just larger' than the controlled set. The results were highly irregular, giving reasonably straight curves (direct ratio) for some of the 'just smaller' sets, but very irregular data for the 'just larger' sets.

San Francisco State College JACK T. TOMLINSON

DISTANCE-JUDGMENTS BY BEES

In his study of bees, Von Frisch has provided us with some unusual psychophysical data—distance-judgments by bees that resemble direct, ratio-estimations by humans.[1] When humans make ratio-estimations of the quantitative aspects of stimuli, the judgments are proportional to a power of the

[1] Karl von Frisch, Gelöste und ungelöste Rätsel der Bienensprache, *Naturwiss.*, 35, 1948, 12-23; *The Dancing Bees*, 1955, 116-126.

magnitude of the stimulus.[2] Does the same rule apply when bees estimate the distance that they have flown from a food supply to the hive? While the human subject assigns numbers to the stimuli, the bee adjusts the speed of a dance with distance. The bee dances (runs forward and turns around) more rapidly after returning to the hive from a nearby food supply than after returning from a distant supply. If the power law applies, the speed (S) of the dance should be proportional to the distance (D): $S = kD^{-c}$. The exponent is negative because the speed decreases as the distance increases. Von Frisch's data do not obey this law. Instead, the speed of the dance is proportional to the logarithm of the distance: $S = -k \log D$. Familiar? It should be. This is Fechner's law. Being good German bees, Von Frisch's subjects obey this ancient dictum. Fig. 1 shows how well the empirical points supplied by

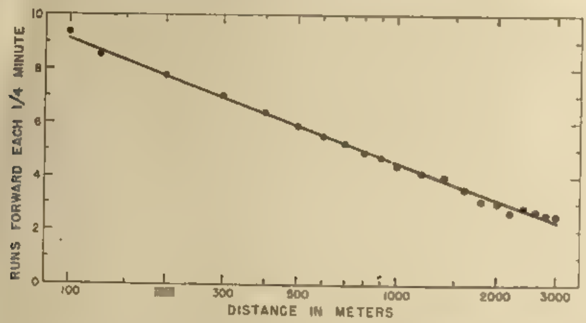


FIG. 1. THE SPEED OF THE BEE'S DANCE AS A FUNCTION OF DISTANCE FLOWN

Von Frisch form a straight line on a semi-logarithmic plot.[3] The ordinate shows the number of times that the bee runs forward each quarter of a minute and the abscissa shows the distance on a logarithmic scale.

Does this mean that Fechner's law holds for lower organisms if not for human beings? We can hardly generalize from this one instance, particularly since the measured stimulus, *i.e.* the distance flown, cannot be the stimulus for the dance.[4] The distances flown are too long ever to act as direct, unitary stimuli. Probably the bee's dance is controlled by some cumulative physiological effect of the flight. Whatever the proximal stimulus may be, it has not been measured. Therefore, the power law is not ruled out for the relation between the speed of the dance and the contemporary stimulus which remains unspecified and unmeasured.

---

[2] S. S. Stevens and E. H. Galanter, Ratio scales and category scales for a dozen perceptual continua, *J. exp. Psychol.*, 54, 1957, 377-409.

[3] von Frisch, *op. cit.*, 1948, 13.

[4] Indeed, the distance is related even less directly to the response than the duration of the flight. After the time of a flight has been lengthened by an unfavorable wind, the bee dances more slowly as if she had flown a greater distance.

At this point, the psychophysicist might ask: How do humans judge a physical variable that is not present as a unitary stimulus at the moment of judging but, instead, is represented by some internal condition that has accumulated over time? The way in which bees react to distance suggests that humans would make direct estimations of similar variables, such as long time-intervals, in accordance with Fechner's law rather than the power law: *i.e.* their judgments would be proportional to the logarithm of the magnitude of the stimulus. Unfortunately such direct judgments have not been made under experimental conditions. It would be interesting to see what the data are like—to see if we can make predictions about human perception from the behavior of lower organisms.

Northeastern University BERTRAM SCHARF

THE 1960 ANNUAL MEETING OF THE AMERICAN PHILOSOPHICAL SOCIETY

The annual meeting of the American Philosophical Society, "held at Philadelphia for promoting useful knowledge," took place on April 21-23, 1960. The psychologist members present were Leonard Carmichael, C. H. Graham, C. I. Hovland, Wolfgang Köhler, W. R. Miles, C. P. Richter, and S. S. Stevens.

The Karl Spencer Lashley Award in the field of Neurobiology was made to Heinrich Klüver, who is Sewell L. Avery Distinguished Service Professor of Biological Psychology at the University of Chicago. A check for $2000 was presented to Dr. Klüver on the occasion of the annual dinner of the Society.

A special session of the meetings was devoted to "current issues and readjustments in American education." In other sessions there were several papers of immediate interest to psychologists. Hadley Cantril discussed the nature of perception, and S. S. Stevens described the evidence for the law that sensation grows as a power-function of the stimulus. Edwin H. Land described and demonstrated the production of full color-effects by the use of only two wavelengths.

Edwin G. Boring will shortly succeed to the chairmanship of the Committee on Publications of the Society.

Harvard University S. S. STEVENS

THE 1960 ANNUAL MEETING OF THE NATIONAL ACADEMY OF SCIENCES

The National Academy of Sciences held its annual meeting at the home of the Academy in Washington, April 25-27, 1960.

Papers were read by J. P. Guilford on "A morphological model for

human intelligence," and Carl Pfaffmann on the "Sensory and motivating properties of taste stimuli." Guilford described a 120-cell orthogonal model of 3 parameters—5 operations × 4 contents × 6 products—which accommodates a great many abilities already revealed by factor-analysis and provides room for more that may be discovered or developed. Pfaffmann described receptor-information for taste as distinguished in afferent impulses in the chorda tympani and the medullary and thalamic relays. He then discussed taste-preferences, positive and negative, as they depend upon the intensity of the exciting stimulus and upon degree of ingestion of the stimulus-material.

There was a symposium on "Current investigations of the brain and behavior," given by four investigators from the National Institute of Mental Health and four from the Walter Reed Army Institute of Research. E. V. Evarts spoke on the "Effects of sleep and waking on single cortical neurones," H. E. Rosvold on the "Relation between cerebrocortical functions and behavior," and Robert Galambos on the "Functional localization of learning." The later papers dealt with the functions of the limbic systems. D. McK. Rioch summed up the discussion, which was presided over by Leonard Carmichael.

The Academy elected to membership, on recommendation of the psychologists, physiologists and zoologists, Robert Galambos of the Walter Reed Army Institute of Research, Donald R. Griffin of Harvard University, Carl I. Hovland of Yale University, and Roger W. Sperry of the California Institute of Technology.

The psychologists in attendance were F. A. Beach, E. G. Boring, Leonard Carmichael, C. H. Graham, J. P. Guilford, H. F. Harlow, Heinrich Klüver, D. B. Lindsley, W. R. Miles, Neal Miller, Carl Pfaffmann, C. P. Richter, and S. S. Stevens.

Harvard University EDWIN G. BORING

FIFTY-SIXTH ANNUAL MEETING OF THE SOCIETY OF EXPERIMENTAL PSYCHOLOGISTS

The fifty-sixth annual meeting of the Society of Experimental Psychologists was held at the University of Wisconsin on April 1 and 2, 1960. W. J. Brogden, Chairman of the Society for the year, served as Chairman of the meeting.

Members present were: Bartley, Bray, Brogden, Estes, Fitts, Geldard, Graham, Grant, Harlow, Humphries, Kappauf, Kennedy, Lawrence, Lindsley, Melton, Morgan, Neff, Newman, Osgood, Pfaffman, Postman, Riggs, Stellar, Underwood and Wickens.

R. C. Davis, Howard H. Kendler and Hans-Lucas Teuber were elected to membership in the Society, bringing the total membership to 61 members and 21 fellows.

Members presented reports of research in progress at scientific sessions held the afternoon of April 1 and the morning of April 2.

The Warren Medal for 1960 was awarded to Carl Pfaffman "for his superb investigations of the discriminatory and motivational functions of taste and smell by a powerful combination of neurophysiological and behavioral techniques."

The Society accepted the invitation of Princeton University to meet in Princeton in 1961. R. M. Gagné was elected Chairman for 1960-61.

Princeton University JOHN L. KENNEDY

FIFTY-SECOND ANNUAL MEETING OF THE SOUTHERN SOCIETY FOR PHILOSOPHY AND PSYCHOLOGY

The Southern Society for Philosophy and Psychology held its fifty-second annual meeting at the Buena Vista Hotel in Biloxi, Mississippi, on April 14-16, 1960. The Departments of Philosophy and Psychology of Florida State University and the University of Florida were hosts.

Twenty papers were presented in philosophy, and twenty-four in psychology. The special award papers were presented by William Thomas Blackstone (University of Florida), on "Can science justify an ethical code?", and by Daniel Smersh Lordahl (Washington University), on "Concept identification as a function of the amount of irrelevant information in simultaneously presented visual and auditory signals." These are selected annually from among those papers presented by members who have not yet received their Ph.D. degrees, or who have received them within the past five years.

At the first joint session the "History of the Southern Society for Philosophy and Psychology" was presented by Marjorie S. Harris (Randolph-Macon Woman's College). Harold N. Lee (Tulane University) served as chairman of the second joint session in which a paper was presented by Donald R. Meyer (Ohio State University) entitled "The central nervous system and the mind-body problem revisited." Carl H. Hamburg (Tulane University) commented on Dr. Meyer's paper.

The presidential address, given by Wilse B. Webb (University of Florida) was entitled "The choice of the problem."

At the annual business meeting 6 philosophers and 26 psychologists were elected to membership. Five philosophers and three psychologists were elected to associate membership.

The following officers were elected for 1960-61:

Rubin Gotesky (University of Georgia), President; William M. Hinton (Washington & Lee University), President-elect; Dan R. Kenshalo (Florida State University), Secretary. Edgar Henderson (Florida State University) is in his third year of a three-year term as Treasurer. James Vanderplas (Washington University), and Leroy Loemker (Emory University) were elected as Council Members from Psychology and Philosophy, respectively.

The Council appointed a Membership Standards Committee, composed of Wilse B. Webb (University of Florida), L. E. Hahn (Washington University), and Stanford Ericksen (Vanderbilt University) to recommend amendments for the Constitution with regard to qualifications for membership. The role of the SSPP is that of making available for psychologists and philosophers an organization devoted primarily to providing a platform and a congenial convention atmosphere for the academic-experimental-theoretical aspects of our discipline, in contrast to those activities generally designated as professional or applied in nature. It was felt that the Constitution should be amended to minimize the possibility of individuals applying for membership under false impressions concerning the nature of the Society.

The membership accepted the invitation of Emory University to hold its next meeting in Atlanta, Georgia.

Florida State University DAN R. KENSHALO

SIXTH ANNUAL MEETING OF THE SOUTHEASTERN PSYCHOLOGICAL ASSOCIATION

The Southeastern Psychological Association held its sixth annual meeting March 31 to April 2nd at the Biltmore Hotel in Atlanta, Georgia. The program consisted of 10 symposia, 71 papers, 2 invited addresses, 1 workshop and 4 special meetings. Approximately 575 people registered for these meetings. The presidential address, "The roles of psychological associations" was presented by M. C. Langhorne. Officers for 1960-61 are: President, J. F. Dashiell; President-Elect, Stanford C. Ericksen; Secretary-Treasurer, Susan W. Gray; Members-at-large of the Executive Committee, Richard W. Husband, Paul Siegel, and Irwin A. Berg. The 1961 meeting will be held in Gatlinburg, Tennessee, April 13 to 15.

George Peabody College for Teachers SUSAN W. GRAY

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Essentials of Psychological Testing. By LEE J. CRONBACH. Second edition. New York, Harper and Brothers, 1960. Pp. xxi, 636. $7.00.

The second edition of Cronbach's *Essentials of Psychological Testing* is among the most teachable introductions to psychometric theory and practice currently available to students of testing. In its much improved format, it compares favorably with Anastasi's *Psychological Testing* for comprehensiveness, and with Thorndike and Hagen's *Measurement and Evaluation in Psychology and Education* for simplicity of presentation.

The revised edition has retained much of the original structure—I. Basic Concepts, II. Tests of Ability, and III. Tests of Typical Performance; each part, however, has been considerably reorganized and expanded. In Part I the exposition of test validity is presented in a separate chapter rather than as one of several considerations in the selection of a test. The controversial concept of construct validity is treated in greater detail. In adopting the concepts and language of decision theory to introduce the purposes of testing and the concepts of validity, Cronbach foreshadows their more extensive application to questions of test utility in selection and classification problems. Besides the usual considerations regarding the need for standardized testing conditions, Cronbach's discussion of the examinee-examiner relationship and test-taking motivations stimulates thinking toward a general theory of test-taking behavior. Finally, from an instructional point of view, continual reference to Bennett's *Test of Mechanical Comprehension* and Kohs' *Block Design* illustrates the discussion on test administration and clarifies that statistical presentation on scoring, reliability, and validity.

Several comments may be made regarding the relative exphasis on certain topics in Part I which should not, however, detract from the general excellence of the introductory material. First, while much space is devoted to specific delineation of the APA code for test distribution, reference to the APA *Technical Recommendations* is rather casual. Secondly, considering Cronbach's statement that "A decision about the merit of a test must come after study of the test manual and accompanying information, the Buros yearbook reviews, and other sources," the treatment of sources of test information seems rather cursory. Lastly, and perhaps out of concern for elementary and secondary school teachers who are likely to study the text, there is a lack of treatment of the mechanics of item-writing and too little on the general test construction process.

Part II contains several noteworthy changes. The chapter on the diagnostic value of the Wechsler profile has been justifiably replaced by greater consideration of the utility of differential aptitude batteries. The material on factor analysis has been brought up-to-date to include, for example, Guilford's theoretical work on the structure on intellect, and, in the chapter on tests of special abilities, Fleishman's praiseworthy studies of psychomotor performance. Finally, the chapter on proficiency tests

is enriched by the discussion of Bloom's *Taxonomy of Educational Objectives.*

Although parts I and II are clearly written and highly stimulating, part III generates and sustains an even higher level of interest. This is due, no doubt, to the intrinsic character of the material, but still more to the excellent quality of the analytic and evaluative commentary. Particularly appealing is the discussion of the bandwidth-fidelity dilemma in the summary chapter appraising clinical and actuarial approaches to the assessment of personality.

In brief, here is an outstanding introduction to current psychological testing. There is no question that the present edition adds to the quality of its almost classic predecessor.

Brooklyn College BENJAMIN ROSNER

Reflections on the Human Venture. By HADLEY CANTRIL and CHARLES H. BUMSTEAD. New York, New York University Press, 1960, Pp. xiv, 344, $6.50.

The attempts to understand "human nature" at the level of scientific conceptualization and at the humanistic level can be profitably combined by what Cantril and Bumstead call the "transactional" view of psychology. Both scientist and humanist can profit by the "holistic" approach to the study of man, the former by an enriched and deepened insight, the latter by a clearer and more precise understanding. With this purpose in mind, the authors have compiled an admirable array of passages from such varying sources as Augustine, Tolstoy, Whitman, Auden, Plato, Paine, Whitehead, and Dewey. Of the 326 pages of text almost 200 are composed of excerpts from philosophers, poets, playwrights, novelists, and scientists. A considerable part of the merit of this book derives from these selections rather than from the text.

The authors' general thesis is that a proper understanding of man must seek to integrate the psychological evidence of the laboratory with the inspiration and insight of the poet. Perception, learning, motivation, the relation of individual to society, etc. may be treated as separable and distinct areas of research only by abstracting them from the on-going, living, human venture. The results of such research must be recombined, with the help of the philosopher, the poet, and the theologian, into a systematic view of human experience.

The particular psychological theories developed in *Reflections on the Human Venture* do not claim originality. Neither, unfortunately, can they claim precision. To examine only one, the authors' reject as inadequate the unguarded extrapolation of the concept of physiological and psychological "homeostasis" as the grounds for a theory of motivation. According to the "transactional view," ". . . man cares more for the process of attaining goals than for the goals themselves." (77) Yet this seems to be modified by the introduction of the notion of "value quality" and "value satisfaction" as if an experience which has these qualities thereby possesses something more than what we ordinarily understand when we label experiences as ". . . thrilling, boring, delightful, disappointing, threatening, and the like." (79) The authors' commit what many now consider the error of conceiving "value" as a quality in the same sense as "red," "thrilling," "sweet," and "disappointing."

In a very curious passage concerning the nature of value *judgments*, it is suggested that the adequacy of a value judgment in so far as its distinctively predictive character is concerned, may be estimated in advance of the performance of the

action which the judgment requires by noting (they say "sensing") the ". . . value-quality you experience in the process of arriving at the judgment." (173) This is perhaps only a slip in the exposition of the theory, for there is no logical relation between the value-quality attaching to the process of judging and the material truth of the judgment made. On the other hand, if the authors mean only to assert that we make predictions about the future on the basis of values realized in past experience, this, of course, is unobjectionable but, as an explanation, trivial. Unfortunately, confusions of this sort are to be found at several key places in the book.

The professional psychologist and philosopher should find this book interesting as an anthology of the better literary expressions of man's experience, while anyone concerned with literature should find it a fascinating and highly readable synthesis of the humanistic and scientific conceptions of "man's venture."

University of Texas D. F. GUSTAFSON

Tests and Measurements: Assessment and Prediction. By JUM C. NUNNALLY, JR. New York, McGraw-Hill, 1959. Pp. viii, 446. $6.95.

Despite the proliferation of books in the field of tests and measurements, this text makes a useful contribution. It is a sophisticated book which offers a solid theoretical framework, an adequate survey of the testing areas, and a valuable evaluation of the future of the field. It is comparable to the standard texts in the area in some ways, although it possesses greater breadth and perhaps less detailed coverage.

As an example of the book's scope, Nunnally is unequivocal regarding the inclusion of statistical concepts. The first third of the book is largely statistical; separate chapters are devoted to the predictive efficiency of tests and to multivariate prediction. Some concepts which fail to find their way into the index of other texts—discriminatory analysis, the standard error of estimate, the Taylor-Russell tables—are discussed comprehensively. This section of the book is often difficult—certainly too difficult for the average undergraduate in an introductory course on testing—but parts may be omitted in such cases.

The standard areas of testing are covered in separate chapters on general intelligence tests, tests of special abilities, achievement tests, opinion-and-attitude measures, and personality measurements. An example of the tests in each area is presented in some detail, but the author avoids the dull recitation of the description of one test after another, which characterizes some books in the field. The general emphasis is put on the uses of the tests and their limitations. Throughout, the necessity of distinguishing between assessment and predictive purposes is re-stated.

The text's wide scope may be thought of as its greatest strength, or as a limiting feature. If the reader uses previous texts as a criterion, he may question the appropriateness of the material on the construction of classroom examinations and the description of recent research techniques. Though the specific selection reveals the author's bias (Q-sort, semantic differential, ASo, etc.), the inclusion of these research techniques gives the reader an understanding of the direction of the field's development.

In summary, Nunnally's text is an integrated presentation of the field of tests and measurements, with a strong theoretical structure. It may suffer from trying to be "all things to all people." The undergraduate students just wanting to know

"what the tests are like" should probably go elsewhere, but the person wanting coverage of the issues in a direct and non-dogmatic way will do well to have the book on his shelf.

George Peabody College for Teachers LAWRENCE S. WRIGHTSMAN, JR.

Constructing Evaluation Instruments. By EDWARD J. FURST. New York, Longmans, Green and Co., 1958. Pp. ix, 334. $4.75.

Constructing Evaluation Instruments is the most recent of a series of books concerning measurement and evaluation published by Longmans, Green and Co. It is also the shortest, a factor which certainly influences its utility as a textbook. In relatively few pages Professor Furst attempts to discuss the principles and applications of measurement primarily as they are present in education and secondarily as they are present in business, industry, government, and the military services. To be sure, this is a difficult task and one in which even a writer as skillful as Professor Furst cannot hope to succeed.

The book is composed of two parts. Part I is devoted to the "really fundamental problems" of evaluation. These problems are (1) determining what to evaluate, (2) identifying the subjects's behavior which is to be observed, (3) selecting suitable situations within which the observations can be made, and (4) recording and summarizing the results of the evaluation effort. Part II is a discussion of the construction of achievement tests. It follows the traditional path of planning the test by means of a table of specifications, constructing test items in terms of the specifications, and administering, scoring, and analyzing the test.

Part I is the longer of the two parts and the more substantial. The author has indeed analyzed some of the basic problems one confronts when constructing an evaluation instrument. In so doing he has avoided a weakness of many books on evaluation, which is an extreme emphasis on the techniques of evaluation with little thought concerning the true nature and purposes of evaluation. Part II is essentially unexciting. More complete and more helpful discussions of achievement test construction are relatively easy to find.

The purposes of this book are to help the reader (1) acquire an understanding of the basic problems involved in developing an instrument for evaluation, (2) increase his knowledge of test construction, (3) gain familiarity with references in evaluation, and (4) broaden his understanding of the purposes served by evaluation instruments. In comparison to other available books on evaluation, this one serves the first and fourth purposes quite well, and the third purpose adequately. Among other factors, space limitations prevent it from achieving the second purpose to anything more than a fair degree.

Cornell University J. STANLEY AHMANN

Eye, Film, and Camera in Color Photography: By RALPH M. EVANS, New York. John Wiley & Sons, 1959. Pp. vi, 410. $8.95.

A fair first approximation to an evaluation of this book is that its psychology should be helpful to photographers and its photography to psychologists. When Evans is on physiological optics, straight physics, or simply color and form in

pictures, he is admirable. Only when he pursues the promise on the dust jacket ("The first major work to apply the psychology of vision to photography"), does a psychological naïvety occasionally crop out. Despite these soft spots he does give considerable firm psychological material that is relevant to photography in general and color photography in particular. Some of this consists of the banalities of visual space and color perception, such as binocular stereopsis, secondary cues, adaptation, contrast, and the like. But he does much more. When, for example, he treats of the eye as a null point instrument, and of the more general problem of tolerance of deviations from the colors of reality under what he terms the consistency principle, he commands convincing research material and a personal capacity for seeing new relationships and making sound generalizations.

However, in the relation between neural activity and perception he has not brought the problem into manageable form. Though in one instance he is sophisticated enough to define "to see" as visual perception, thereafter he consistently refers to the eye or brain as seeing or having perceptions, apparently unaware that these organs merely provide the preconditions for seeing. He makes the curbstone error (which physiologists also often make) of assuming that because the retinal image is in the eye and the ensuing neural processes are in it and in the brain the perception must be in these organs. Are not the obvious grand discrepancies and qualitative alienation between neural activity and ensuing perception a truism for all modalities? This is worth rigorous scrutiny, because in psychology it is precisely here that one begins to tell the men from the boys.

Among the other soft spots is a more tangible one, and strange for a student touched by *Gestalt*. He confuses the necessary dependence of the perception of objects on the composition of the illuminant with the supposed necessity of being conscious of the composition of the illuminant. Such consciousness is not characteristic at all of ordinary seeing. But an expert may train himself to analyze out the illuminant, just as he may discipline himself to isolate a partial from a fundamental or taste from smell in food.

Evans grazes the work of *Gestalters* and para*Gestalters* like Gelb, Katz, and Helson. His timid handling of the constancies, for instance, suggest that these have not drawn out his best talents, and that he has not grasped their fuller implications or their very serious limitations and vulnerabilities. Nevertheless, in handling them at all, he, like many other physicists, accepts the possibility of veering away from crude mechanism, in contrast to the atavism of the psychologists who, having once known better, now lust after regression to crude Watsonian mechanism camouflaged with statistics and analyses of fragments of rats' learning curves.

Like Maugham, who considers it the first object of fiction to please, so Evans, I believe, takes it as the first duty of color photography to please. To this end he marshals his considerable resources from physics, physiology, projective geometry, psychology, and aesthetics. For the psychologist interested in more than pleasing color pictures and with a bent for optics and the physics and physiology of color this book will broaden the foundations of his knowledge of visual perception. But one basic feature is missing, an explanation of the I.C.I. standard data and the quantitative methods of colorimetry developed at M.I.T. and at the Imperial College of Science and Technology. For this material one has to go to Evans' 1948 book

on color. Such powerful quantitative tools are much too universal to leave to the monopoly of physicists and color technologists. They should supplant the color pyramid and its likes—sad things that have survived far beyond their time.

Sarasota, Florida ELLIS FREEMAN

Statistical Analysis in Psychology and Education. By GEORGE A. FERGUSON. New York, McGraw-Hill Book Company, Inc., 1959. Pp. vii, 347. $7.00.

This book is designed as an introduction to elementary statistical treatment of data in psychology, education and related research fields. The author includes an introductory chapter on the basic ideas in statistics, and each chapter contains an introductory section which will be helpful to the newcomer to the statistical world. The scope of the book is sufficiently broad to cover the needs of those who desire a knowledge of statistics for the purpose of reading material which presents data and research results in statistical form, as well as, for most general purposes, those who must design and report research utilizing statistical methods and procedures.

The author begins with the simplest idea of statistical notation and proceeds step by step until his material includes rank order and other correlation methods, transformations, analysis of variance for one and two variables, selected non-parametric tests, and some multivariate problems. This is an ambitious coverage for an introductory textbook of this size, yet present developments in research and writing indicate that this is precisely the coverage needed. Too much time is spent at present by the average student learning to handle simple data by paper and pencil methods, whereas the student needs to gain rapidly a better understanding of statistical methods. The book does not slight basics, for example, it includes a chapter on probability, but before the halfway mark, one is meeting tests of significance and chi square. Most commendable is the presentation of two chapters on analysis of variance, the second giving a full treatment of the two-way classification case in all aspects.

The chief value of the book is that while it presents nothing new in the field of statistics, it covers in a clear style more material in less space than most other books available.

University of Texas G. H. SLUSSER

Readings in the Psychology of Adjustment. Edited by LEON GORLOW and WALTER KATKOVSKY. New York, McGraw-Hill, 1959. Pp. 541. $6.50.

This first-rate book will give the student a handy introduction to the chief trends in a field where a wide range of reading would seem to be more desirable than in almost any other undergraduate courses in psychology departments. After three selections indicating how the general philosophy of science can embrace the behavioral sciences and four illustrations of how scientific approaches work in real studies, there are sections on what normality and mental health are, and on theories of personality, with individual passages giving fundamental positions of Freud, Jung, Adler, Sullivan, Rogers, Maslow, Eysenck, and Shaffer and Shoben.

Under determinants of adjustment are readings on childhood, culture, social class, and family; under dynamics, readings on conflict, anxiety, neurosis, character disorders. Therapy is less thoroughly covered, but this is an undergraduate text and enough is given to suggest to the student what therapy attempts. Under current

pressing problems the authors have reprinted the discussion, from *Science,* by B. F. Skinner and C. R. Rogers on the control of human behavior, an article on the significance of the Durham case for interpreting the responsibility of the individual's conduct, and one on fake academic degrees used by charlatans in the field of mental hygiene.

In a section by E. Fromm there is quoted, "Like the handbag, one has to be in fashion on the personality market. . . . This knowledge is transmitted through education. . . ." And, of course, one has to be in fashion on the theory-of-personality market. This book is fully up-to-date and the best insights from the best thinkers on adjustment are here adequately and tastefully sampled. I should guess that many a student will be tempted to go to the fuller exposition of many of the authors sampled. I would only wish for two things: A little biographical note on each author and a little more economical price—for a supplementary readings book.

U. S. Navy Medical Neuropsychiatric Research Unit, San Diego
WALTER L. WILKINS

Personality Development and Assessment, second edition. CHARLES M. HARSH and H. G. SCHRICKEL, New York, The Ronald Press, 1959. Pp. 536.

The first edition of this unique book appeared ten years ago. The new edition has the same basic structure but has much new writing, more polished chapter titles, has been reshuffled somewhat and also comes to some terms with the massive literature of the decade.

The text is useful for introductory courses in personality where the instructor wants the students to obtain a wide array of concepts, facts, data and methods. The authors still relentlessly avoid putting forward a systematic theory of their own, on the basis of the reasonable assumption that the uniqueness and richness of personality is not yet encompassed by any existing theory and method. Generally the patience with the data and thinking of the field is retained, as is the consistent challenge to early "closure." It is hopeful to notice, however, that new evidence and ten years has led at least one of the authors to be a bit more explicit in outlining his own scheme for integration in chapters 4 and 5.

The uniqueness of this text lies in how it tries to cover the field. The book begins with a developmental emphasis in the context of a scheme of the analytic elements necessary to understanding performance as well as change. There is fresh emphasis on notions of "evolving concepts of the self" here. The portrait then continues through to old age. Having stated the problems in these first twelve chapters, the authors then turn to a critical survey of theories. Field theory is more explicitly considered than before, along with the psychoanalytic movement and other traditional topics. A sensitive and critical discussion of techniques of measurement and of modes of interpreting data follows. Next comes a good introductory discussion of factor analytic work on organization of personality, and then a discussion of personality integration which also helps to integrate the volume.

This venture is an example of collaboration in which the individual authors have retained their own identity, style and bold expression of opinion. This zestful property of the book has been heightened. A good text for the teacher who wishes to introduce his students to the wide array of notions in the area even while choosing to lecture from a more systematic position.

W. W. LAMBERT
Cornell University

Bodily Structure and the Will. By M. A. MACCONAIL. London, The Aquin Press, 1960. Pp. 28. 2s.

MacConail's monograph is Paper No. 34, one of a series of 35 delivered as lectures to the Aquinas Society of London. The author, a somatologist, teaches in the Department of Anatomy at University College, Cork, Ireland. His purpose is to show that the notion of freewill is not opposed to contemporary knowledge of bodily structure and behavior, nor these latter opposed to it. This assertion is made in reply to the increasingly vigorous attacks from certain scientific quarters upon the notion of the composition of the human person from a material and a spiritual part. This paper carries the imprimatur of the Roman Catholic Church and MacConail is an avowed Thomist.

The plan of the paper is as follows. First, the author considers the idea of a reciprocal causal interaction of a human body and a human soul, attempting to show that this relationship is the one suggested by the Canon of Parsimony. Secondly, he applies the same canon to show that, in whatever measure body and soul are distinguishable, the channel of action of each upon the other is the same in the field of muscular activity. Finally, he shows that the notion of freewill is not incompatible with contemporary knowledge of bodily structure.

The argument proceeds using the tools of symbolic logic and, although the author presents carefully his matrices of operational relational statements, the average reader, not acquainted with recent symbolic logic, will not be enlightened by his use of "proto-Boolean algebra." Fortunately, at the close of the paper, there is provided a brief summary of the argument in ordinary language.

University of Texas G. H. SLUSSER

Longitudinal Studies of Child Personality. By ALAN A. STONE, M.D., and GLORIA COCHRANE ONQÚE, M.D. Published for The Commonwealth Fund by Harvard University Press, Cambridge, Massachusetts, 1959. Pp. xiv, 314. $5.00.

The authors of Longitudinal Studies of Child Personality, Drs. Stone and Onqúe, had begun working as a medical students, under the direction of Dr. Milton J. E. Senn, Director of the Yale Child Study Center, on a project utilizing the longitudinal approach. They, among a group of medical students, began abstracting longitudinal studies in the literature and after the group effort had lapsed, they obtained permission from the Yale faculty to complete and to present the annotated bibliography as their thesis.

The present volume is based upon this thesis. It contains 297 of the some 1000 articles and books published through 1955 that they had examined. Their criteria for inclusion were: (1) that the study be truly longitudinal in nature *i.e.* be a "forward-looking approach which aims to study the individual as he passes from one stage to a later, and to gauge the influence of succeeding experiences and circumstances" (p. xi); and (2) that the study be concerned only with emotional and social behavior in infants and children.

The authors admit that it was difficult at times to decide what clearly fell within their conception of the longitudinal approach. They critically analyzed this approach by stating that longitudinal studies invariably are faced with the problems of "(1) small samples, (2) poor testing devices, and (3) inability to convert theoretic hypo-

thesis into experimental or operational terms which will define the significant variables" (p. xii).

Their abstracts, which form the body of the book, follow a set outline: (1) the bibliographical reference; (2) the setting for the study; (3) subjects used; (4) the time-span, methods of observation and testing; (5) the findings and the interpretations of the author of the report. Very seldom did the authors vary from this approach, although, as would be expected, some of the studies could not neatly be condensed into that framework.

Although this volume will give students of child development a cursory view of the research in this field, the brevity and style of the abstracts unfortunately allow very little basis upon which to view the studies critically. It will be very difficult, therefore, for the beginning student in particular to separate the methodologically sound reports from the less adequate ones. It is also unfortunate that the authors did not choose to integrate in a final chapter the present day trends which they saw emerging from their review of the literature. By so doing, they would have added a creative touch to the work.

University of Houston — RICHARD I. EVANS

Psychology in Business. By LESLIE BEACH and ELON L. CLARK. New York, McGraw-Hill, 1959. Pp. ix, 313. $4.00.

Teachers faced with the task of selecting a readable, informative and scientifically sound text for college students or managerial personnel who have not taken an introductory course in psychology will do well to examine this new book. Simply and clearly written, the material lends itself to a ready understanding of most of the facts and principles of psychology needed by those who seek to apply psychology to business and work relationships.

The first part of the book deals with fundamental concepts in the areas of individual differences, intelligence, learning and rembering, emotion, motivation and frustration, and perception; the second with applications of basic principles to problems in business such as selection, training, evaluation, promotion, and supervision of employees, as well as customer and community relations. Discussion of these content areas is in an informal straightforward style with very few direct references to the supporting literature. Much of the material has been class-tested with students preparing for business careers at General Motors Institute.

Although the book was designed primarily for students who have not had previous courses in psychology, it could serve as a basic text for advanced courses in psychology, it could serve as a basic text for advanced courses in industrial or business psychology. Supplementary reading material would then be in order since the text covers material usually included in the first course. This reviewer's experience in teaching advanced courses in industrial psychology leads him to conclude that a review of basic principles is necessary or at least very helpful if the student is to realize fully the application of psychology to industrial and business problems. Unfortunately, students who have had the introductory course seldom see the basic principles in the perspective needed for application to industrial and business situations.

The chief negative criticism that can be levelled against the book is the inadequate treatment of the nature of the business organization and the problems that

arise in large organizations. The pyramidal shape of business organizations, communication problems, the organizational atmosphere and organizational theory are topic areas that are omitted or briefly mentioned, but which merit consideration in a book which purports to examine psychology at work in business and industry.

Carnegie Institute of Technology HARRY W. KARN

Orthopsychiatry and the School. Edited by MORRIS KRUGMAN. New York, American Orthopsychiatric Association, 1958. Pp. x, 265.

The nature of the volume is presented well by Dr. Krugman in his introduction, "In publishing these twenty-six articles, the purpose was not to present a unified point of view, or an officially sanctioned approach, but rather to portray a wide variety of approaches in differnet situations so that the reader will be in a position to weigh and evaluate different programs and to draw from them suggestions and practices that may be applicable or useful for a specific situation." (p. x) In essence the book suggests the "newness" or "immaturity" of the field. The generalizations that one can draw are also well illustrated by Krugman. "One can find support for the thesis . . . that teachers are poor diagnosticians, and that teachers are superior diagnosticians; that educators do not make good therapists, and that educators are naturally very good therapists . . ." (pp. x and xi)

This state of affairs wherein every generalization is "true" even though it contradicts another one or two "true" generalizations will not make the experimental psychologists nor the research oriented educator happy. To plead recency in the interrelationships of psychiatry and education approaching problems together is to rationalize or indicate ignorance. Adolph Meyer wrote in 1908, "Mental Hygiene has made decided strides in our schools during the last decade. . . . It would probably be wrong and certainly utopian, to expect that schools should create special classes for every special emergency. I rather look forward to the establishment of school hopsitals or hopsital schools, which will undertake at the proper time the investigation of special difficulties presented by certain cases, and provide for the training of the child and the parents." Contradictory opinions and testimonial evidence have existed for at least fifty years. The reviewer is concerned in that this collection of papers continues in this tradition rather than carefully specifying areas of fundamental conflict and moving toward a careful solution of the immensely complex theoretical and practical problems involved in orthopsychiatry and the school.

The book's strength, in the reviewer's opinion, lies in the insights produced by careful experienced observers in the area. The volume would make an excellent base for a series of carefully controlled studies of the impact of the school: teacher attitudes and interaction, group norms and group structure, and patterns of activities and programs on the mental health of children. Similarly, the sociologically oriented educator will find exciting hypotheses in the penetrating comments—repeated throughout the volume—concerning the interrelations of professional roles, conflicting expectations, and overlapping skills and training that exist within the area of mental hygiene services to the school age population.

The book contains few references and bibliography for the student; there is no index.

Washington Univ. LOUIS M. SMITH

Readings in Social Psychology. Edited by Eleanor E. Maccoby, T. M. Newcomb, and E. L. Hartley, Third Edition. Prepared for the Committee on the Teaching of Social Psychology of the Society for the Psychological Study of Social Issues (SPSSI). New York, Henry Holt and Company, 1958. Pp. xi, 674. $6.75.

Few works in social psychology have had the persistent success or widespread notice achieved by this collected work, now revised again in a third edition. As a teaching aid supplementing textual material it has found a valued place. But in another more significant sense it serves in part to shape the field. Its periodic revision creates both an impression and redefinition of social psychology for a wide audience; and, to some within the field, the selection of content might seem to encourage or discourage particular lines of activity. The responsibilities of editorship are therefore bound to be great and likely to yield a diversity of view. Of the 62 papers in the current edition, considerably more are drawn from the immediate past. Fewer of the older familiar papers are in evidence, notably ethnological works and discursive-theoretical contributions. Contemporary empiricism, particularly along experimental lines, receives greater emphasis than before; one therefore notes a corresponding reduction in works conveying the flavor of tradition and theory. Some compensation is afforded by several specially prepared and quite good commentaries —Bruner on perception in social psychology, Bronfenbrenner on socialization, Bales on roles in problem-solving groups—which provide a useful integration of previous work and thought in an area. Withal, the collection sustains interest. The volume continues to justify itself as a resource for the student at any level.

Washington University — E. P. Hollander

The Psychology of Exceptional Children. By Karl C. Garrison and Dewey G. Force, Jr. Third edition. New York, The Ronald Press Company, 1959. Pp. vi, 586, $6.00.

This refreshing revision is a significant improvement of what has become a standard text and reference work for classroom teachers of atypical school children. First published in 1940 it quickly became a popular source of readily available material in the preparation of prospective teachers or special education. A revised edition appeared in 1950 with appreciable reorganization and expansion. Published now in co-authorship and fully modernized the new material has been assimilated without pages of extension but with reorganization which reveals the movements it portrays.

The emphasis is still on psychological knowledge relevant to special education and the role of the classroom teacher in relation to the welfare of exceptional children. There are new chapters on epilepsy, cerebral palsy and cardiac conditions. The chapter organization now includes neurological with health impairments. The earlier appendices on test and record forms and book lists for supplementary readings are replaced by a technical glossary and a directory of motion pictures.

Readers seeking scholarly erudition or research compendia may prefer other sources. But the new student will thank the authors for selecting and interpreting substantial and relatively recent studies for him. The treatment is earnest rather than patronizing and successfully reflects the marked advances in this field since this book was first launched two decades ago.

Bellingham, Washington — Edgar A. Doll

BOOKS RECENTLY RECEIVED FOR REVIEW

Adler, Alfred, *The Education of the Individual*, New York, Philosophical Library, 1958, pp. vii, 143, $3.50.

Albee, George W., *Mental Health Manpower Trends*, Joint Commission on Mental Illness and Health, Monogr. Ser. No. 3. New York, Basic Books, Inc., 1959, pp. v, 361, $6.75.

Balint, Michael, *Thrills and Regressions*, New York, International Universities Press, 1959, pp. 148, $4.00.

Bergler, Edmund, *Principles of Self-Damage*, New York, Philosophical Library, 1959, pp. v, 469, $6.00.

Berkowitz, Pearl H. and Esther P. Rothman, *The Disturbed Child: Recognition and Psychoeducational Therapy in the Classroom*, New York, New York University Press, 1960, pp. vii, 204, $4.00.

Bickel, Lothar, *The Unity of Body and Mind*, New York, Philosophical Library, 1959, pp. 167, $3.75.

Chance, Erika, *Families in Treatment: From the Viewpoint of the Patient, the Clinician and the Researcher*, New York, Basic Books, Inc., 1959, pp. v, 234, $5.50.

Da Vinci, Leonardo, *Philosophical Diary*, New York, Philosophical Library, 1959, pp. 87, $2.75.

Dewey, John, *Dictionary of Education*, (Edited by Ralph B. Winn), New York, Philosophical Library, 1959, pp. v, 150, $3.75.

Dewey, John, *Moral Principles in Education*, New York. Philosophical Library, 1959, pp. v, 61, $2.75.

Dorsey, John M. and Walter H. Seegers, *Living Consciously: The Science of Self*, Detroit, Wayne State University Press, 1959, pp. vii, 174, $4.95.

Drake, L. E. and E. R. Oetting, *An MMPI Codebook for Counselors*, Minneapolis, University of Minnesota Press, 1959, pp. 140, $3.75.

Dunham, H. Warren, *Sociological Theory and Mental Disorder*, Detroit, Wayne State University Press, 1959, pp. ix, 298, $5.50.

Eliasberg, Wladimir G., *Psychotherapy and Society: Psychotherapy for the Many and the Few*. New York, Philosophical Library, 1959, pp. vii, 223, $6.00.

Feifel, Herman, *The Meaning of Death*, New York, McGraw-Hill, 1959, pp. v. 351, $6.50.

Friedenberg, Edgar Z., *The Vanishing Adolescent*, Boston, Beacon Press, 1959, pp. ix, 144, $2.95.

Fromm-Reichmann, Frieda, *Psychoanalysis and Psychotherapy*, Selected Papers, Edited by D. M. Bullard, Chicago, The University of Chicago Press, 1959, pp. v, 350, $7.50.

Frosch, John, and Nathaniel Ross, *The Annual Survey of Psychoanalysis: A Comprehensive Survey of Current Psycoanalytic Theory and Practice*, Volume V, New York, International Universities Press, 1959, pp. xi, 608, $12.00.

Grenander, Ulf (ed.), *Probability and Statistics*. The Harold Cramér Volume, Stockholm, Almquist & Wiksell, and New York, John Wiley & Sons, 1959, pp. 434, $12.50.

Hall, Edward T., *The Silent Language*, Garden City, Doubleday & Company, 1959, pp. 240, $3.95.

Hartley, Eugene L. and Gerhart D. Wiebe, *Casebook in Social Processes,* New York, Thomas Y. Crowell Co., 1960, pp. ix, 534, $3.75.

Honigmann, John, J., *The World of Man,* New York, Harper & Brothers, 1959, pp. vii, 971, $7.50.

Kemeny, John G. and J. Laurie Snell, *Finite Markov Chains,* Princeton, D. Van-Nostrand Company, 1960, pp. viii, 210, $5.00.

Kruger, Judith, *My Fight for Sanity,* Philadelphia, Chilton Company, 1959, pp. 244, $4.95.

Longmore, Laura, *The Dispossessed. A Study of the Sex-Life of Bantu Women in Urban Areas in and around Johannesburg,* London, Jonathan Cape, 1959, pp. 334, 30 s.

Mandel, M., *As the Twig Is Bent,* New York, Jonathan David, 1959, pp. xi, 236, $3.95.

Masor, Nathan, *The New Psychiatry,* New York, Philosophical Library, 1959, pp. vii, 154, $3.75.

Opler, Marvin K. (ed.), *Culture and Mental Health: Cross-Cultural Studies,* New York, The Macmillan Company, 1959, pp. vii, 533, $8.75.

Ostrovsky, Everett S., *Father to the Child: Case Studies of the Experiences of a Male Teacher with Young Children,* New York, G. P. Putnam's Sons, 1959, pp. ix, 173, $3.75.

Palmer, L. S., *Man's Journey Through Time: A First Step in Physical and Cultural Anthropochronology,* New York, Philosophical Library, 1959, pp. vii, 184, $10.00

Rasey, Marie I. (ed.), *The Nature of Being Human. The Franklin Lectures of 1956-1957,* Detroit, Wayne State University Press, 1959, pp. vii, 115, $3.95.

Redemeier, Harry and Jackson Toby, *Social Problems in America: Costs and Casualties in an Acquisitive Society,* New York, John Wiley, 1960, pp. xvi, 510, $6.75.

Rohrer, John H. and Munro S. Edmonson, (eds.), *The Eighth Generation: Cultures and Personalities of New Orleans Negroes,* New York, Harper & Brothers, 1960, pp. xi, 346, $6.00.

Runes, Dagobert D., *A Dictionary of Thought: From My Writings and From My Evenings,* New York, Philosophical Library, 1959, pp. 152, $5.00.

Schachtel, Ernest G., *Metamorphosis: On the Development of Affect, Perception, Attention, and Memory,* New York, Basic Books, Inc., 1959, pp. vii, 344, $6.00.

Siegel, Bernard J., (ed.), *Biennial Review of Anthropology, 1959,* Stanford, Stanford University Press, 1959, pp. vi, 273, $6.00.

Weijel, J. A., *Psychiatry in General Practice,* Amsterdam, Elsevier Publishing Company, 1958, pp. 208, $7.00.

Weinberg, Harry L., *Levels of Knowing and Existence: Studies in General Semantics,* New York, Harper & Brothers, 1959, pp. vii, 274, $3.25.

Wheeler, Wayne (ed.), *Social Change and Mental Health,* Parkville, The Park College Press, 1960, pp. 49, $1.00.

Winter, Edward H., *Beyond the Mountains of the Moon: The Lives of Four Africans,* Urbana, The University of Illinois Press, 1959, pp. v, 276, $5.50.

Zietz, Dorothy, *Child Welfare: Principles and Methods,* New York, John Wiley & Sons, 1959, pp. vii, 384, $5.50.

Zuver, Dudley, *Retreat to Reality,* New York, Thomas Y. Crowell Company, 1959, pp. 281, $4.75.

NEW EDITIONS

Abt, Lawrence E. and L. Bellak, *Projective Psychology: Clinical Approaches to the Total Personality,* Second Edition, New York, Grove Press, 1959, pp. vii, 485, xiv, $5.00 (cloth) or $2.95 (soft cover).

Frazer, Sir James George, *The New Golden Bough, A New Abridgment,* Edited by T. H. Gaster, New York, Criterion Books, 1959, pp. v, 738, $8.50.

Levens, A. S., *Nomography,* Second Edition, New York, John Wiley & Sons, 1959, pp. v, 296, $8.50.

McKinney, Fred, *Psychology of Personal Adjustment: Students' Introduction to Mental Hygiene,* Third Edition, New York, John Wiley & Sons, 1960, pp. xiii, 490.

SOFT COVERED REPRINTS

Beunenkamp, Cornelius, *Fortunate Strangers: An Experience in Group Psychotherapy.* New York, Grove Press, 1959, pp. 269, $1.95.

Hoffman, Frederick J., *Freudianism and the Literary Mind,* New York, Grove Press, 1959, pp. vii, 350, $2.45.

Jones, Ernest, *On the Nightmare* (Reprint of Second Edition), New York, Grove Press, 1959, pp. 374, $1.95.

McCary, J. L. (ed.), *Psychology of Personality: Six Modern Approaches,* New York, Grove Press, 1956, pp. ix, 383, $1.95.

Stekel, Wilhelm, *Patterns of Psychosexual Infantilism,* New York, Grove Press, 1959, pp. vii, 412, $1.95.

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXXIII SEPTEMBER, 1960 No. 3

THE PSYCHOPHYSICS OF FORM: REVERSIBLE-PERSPECTIVE DRAWINGS OF SPATIAL OBJECTS

By JULIAN HOCHBERG and VIRGINIA BROOKS, Cornell University

One of the primary theoretical and methodological problems in psychology is posed by form; as a datum to be 'explained,' as a relationship to be analyzed into elementary components, and as a variable to be manipulated.[1] The earliest treatments of *form* (and most subsequent ones) were concerned with *representational form* (*e.g.* drawings), and the area remains one of considerable theoretical and practical importance.[2]

The psychophysics of *represented form* has not, however, been systematically pursued until very recently, despite the armchair-experiments perpetuated in advanced as well as in introductory textbooks of psychology and art. Past theoretical treatments tended to be more philosophical than psychological in nature.[3] Empiricists have 'explained' pictorial forms as the obvious products of learning, Gestaltist psychologists have 'explained' them in terms of undefined 'figural goodness' and less-defined speculative brain-processes, and others have classed them as ambiguous 'ghost shapes.'[4] De-

* Received for publication August 10, 1959. This paper reports part of series of investigations supported by a grant G 44581 from the National Science Foundation.

[1] Kurt Koffka, *Principles of Gestalt Psychology*, 1935, 129-148, 171-176; E. G. Boring, *Sensation and Perception in the History of Experimental Psychology*, 1942, 221-225, 233-256; E. R. Hilgard, *Theories of Learning*, 2nd ed., 1956, 226-249.

[2] H. W. Hake, Contributions of psychology to the study of pattern vision, USAF WADC Tech. Rep. WADC-TR-57-621, Oct. 1957, 59-98; T. W. Wulfeck and J. H. Taylor, (ed.), *Form Discrimination as Related to Military Problems*, NAS-NRD publication, Nr. 561, Apr. 1957, 3-37, 109-222; Fred Attneave and M. D. Arnoult, The quantitative study of shape and pattern recognition, *Psychol. Bull.*, 53, 1956, 452-471; P. M. Fitts, Meyer Weinstein, Maurice Rappaport, Nancy Anderson and J. A. Leonard, Stimulus correlates of visual pattern recognition: A probability approach, *J. exp. Psychol.*, 51, 1956, 1-11.

[3] Julian Hochberg, Spatial representation: Theme 10, *Proc. Int. Congr. Psychol.*, Brussels, 1957.

[4] J. J. Gibson, What is a form?, *Psychol. Rev.*, 58, 1951, 403-412.

spite their popularity as 'explanations,' neither 'past experience' nor 'cortical organization' is of much use in predicting what will be perceived (or reported) in response to specific stimulation, while a declaration of 'indeterminacy' maintains, in essence, that there is nothing to be predicted where 'ambiguity' prevails. In apparent defiance of this point of view, the stimuli used in the present investigation were reversible-perspective figures—highly 'ambiguous' stimulus-patterns. We shall see, nevertheless, that very specific *lawful* predictions *can* be made in this area, and that there exists a body of human response-regularity which demands systematic study regardless of theoretical precommitments and which provides a sounder basis for theory-formation than we have had before.

The selection of subject-matter was not arbitrary. Reversible-perspective figures (and other examples of ambiguous or 'illusory' depth) have assumed critical importance at several points, on both sides of the theoretical controversy between empiricism and nativism. The fact that one and the same stimulus-figure would give rise to quite different (and convincing) experiences of depth can be taken as evidence against the existence of any fixed, innate mechanisms tying spatial sensations to specific receptor-processes.[5] On the other hand, the identical phenomena also have provided major support for the *Gestalt* claim that space-perception is determined, not by past experience, but by the 'laws' of cortical organization.[6] Furthermore, 'ambiguity' of psychophysical relationship—of predicting response from stimulation—has been made a central issue in the two opposed approaches of 'stimulus-determinism' (typified by Gibson,[7] to whom perception is unequivocally determined by stimulation, if only we choose the right variables of stimulation to examine) and 'probabilistic functionalism' (epitomized by Brunswik,[8] who maintained that we must abandon systematic psychophysics in favor of ecologically-representative functional surveys, because of the purported equivocal relationship between stimulation and perception). More popularly, 'New Look' hypotheses have identified 'ambiguity' with non-stimulus-determined projection of motives, of training, and so forth. We hold here that none of these points of view will help us predict what will be perceived in any specific situation. These issues comprise pseudo-theories with respect (at least) to the present area, since they may lead to 'feelings of understanding' and generate controversy, but do not provide verifiable, predictive knowledge.

This paper reports the development and cross-validation of a (higher order)[9] psychophysical equation which predicts the depth-responses of un-

[5] Hermann von Helmholtz, *Physiological Optics*, J. P. C. Southall, ed., Vol. 3, 1925, 13, 293.

[6] Koffka, *op. cit.*, 159-164.

[7] Gibson, *The Perception of the Visual World*, 1950, 59-76.

[8] Egon Brunswik, *Perception and the Representative Design of Psychological Experiments*, 1956, 26-55.

[9] Gibson, Social perception and the psychology of perceptual learning, in Muzafer Sherif and M. O. Wilson, eds., *Group Relations at the Crossroads*, 1953, 131-136; Hochberg, Effects of the Gestalt revolution: The Cornell symposium on perception, *Psychol. Rev.*, 64, 1957, 73-84.

trained *O*s to non-perspective projection-drawings of three-dimensional objects ('wire forms'). It attempts to substitute objectively measurable variables for intuitive Gestalt concepts ('goodness,' 'simplicity,' and 'organization'), and is completely neutral to such issues as 'nature *vs.* nurture,' 'organization *vs.* atomism,' 'stimulus-determinism *vs.* probabilistic functionalism.'

Experiment I. Determining the Psychophysical Equations

Problem. The specific 'law' tested in the present study was independently proposed by Hochberg and McAlister and by Attneave, in very similar form.[10] The stimulus-figures employed to test it, and the general intent of the procedures, are derived from Kopfermann, who suggested that, if a figure was 'better' as a three-dimensional shape than as a two-dimensional organization, it would be seen as three-dimensional (and vice versa).[11] This proposition is recast here into measurable terms as follows: The more information required to specify each alternate shape which may be reported in response to a given stimulus-figure, the less likely it is that that shape will be seen or reported relative to its alternates. The procedures employed grew out of earlier studies by Hochberg and McAlister.[12]

In the first two experiments reported here, the stimulus-figures were divided into *families* of four to six different reversible-perspective projections or 'views' of a given three-dimensional object. Since all comparisons of relative three-dimensionality were made between the members of each family, only the two-dimensional stimulus-characteristics need here be considered in predicting relative apparent solidity.

This procedure was adopted for three reasons. (1) It permitted a flexible sampling of shapes. (2) It permitted us to consider only the two-dimensional complexity of each stimulus-figure (as a first approximation), since the same three-dimensional complexity must obtain for all members of a correctly-designed family. (Some families were later found to be incorrectly designed, *i.e.* in Plate I, Figs. 21–40 do not share the same tridimensional versions within each family and some are not in reversible perspective.) (3) Previous research by McAlister,[13] and preliminary investigations by the present writers, suggested that it would be much

[10] Julian Hochberg and Edward McAlister, A quantitative approach to figural 'goodness,' *J. exp Psychol.*, 46, 1953, 361-364; Fred Attneave, Some informational aspects of visual perception, Psychol. Rev., 61, 1954, 183-193.

[11] Hertha Kopfermann, Psychologische Untersuchungen über die Wirkung zweidimensionaler Darstellungen körperlicher Gebilde, *Psychol. Forsch.*, 13, 1930, 293-364.

[12] Hochberg and McAlister, *op. cit.* 1953, 361-364; Relative size vs. familiar size in the perception of represented depth, this Journal, 68, 1955, 295-296.

[13] McAlister, An experimental investigation of the relationship between the geometric components and scaled judgments of ambiguous figures. Unpublished Doctoral dissertation, Cornell University, 1955, 37-72.

easier to find a single equation which would predict within families than to find one to encompass cross-family comparisons as well. Since this procedure could quite conceivably decrease the generalizability of both the obtained response-scales, and of the psychophysical equations based upon them, the effects of this (and other) constraints have been specifically examined (*vide infra*), and have been found not to impair the generality of the findings.

Hypothesis. The relative apparent tridimensionality of each member of a family of reversible-perspective representations of a given three-dimensional object will be a simple function of the geometrical complexity of the two-dimensional figure (as a first approximation).

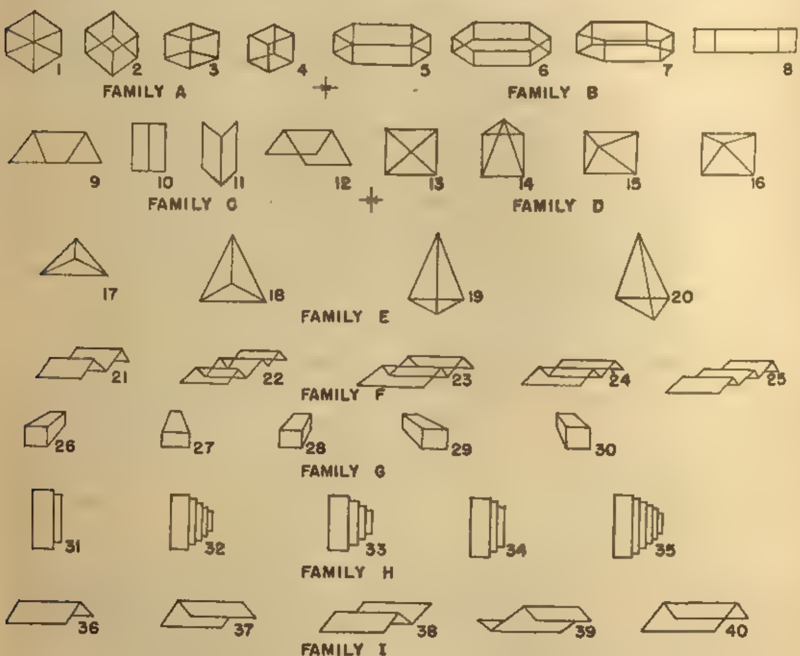


PLATE I. STIMULUS-FIGURES USED IN EXPERIMENT I

Method. The *O*s were college students. The total numbers viewing each family of stimulus-figures were: 378 for Figs. 1–20 and 52 for Figs. 21–40. All *O*s were naïve as to the purposes of the experiment.

Test-booklets were prepared containing (a) instructions; (b) nine families of reversible-perspective figures, one family per page (see Figs. 1–40, Plate I); and (c) a horizontal, numbered scale on each page, running from numbered ends, 0 to 10.

Booklets were administered to groups of *O*s in balanced order of presentation

between and within families of figures. The *O*s were instructed to place the figure which seemed most strongly 'solid' or 'three-dimensional' at 10 on the scale, the 'least solid' or 'most flat' at 0, and then to indicate where the other figures belonged

TABLE I

VALUES GENERATED BY EQUATIONS [1]—TEST 4; [2]—MULTIPLE = REGRESSION; AND [3]—BASED ON FACTOR ANALYSIS

Stimulus		Equation			Test	Stimulus		Equation			Test
Family	Figure	[1]	[2]	[3]	1 obtained	Family	Figure	[1]	[2]	[3]	1 obtained
	1	0.0	0.4	0.0	0.0		36	0.0	−0.2	1.6	2.2
	2	10.0	8.5	9.0	7.5		37	2.5	2.8	5.2	2.8
A	3	6.7	6.3	4.7	5.3	I	38	10.0	7.8	7.5	10.0
	4	10.0	10.0	10.0	10.0		39	10.0	11.0	8.3	10.0
	5	7.5	5.2	4.8	3.7		40	0.0	0.9	4.9	0.0
	6	10.0	9.6	4.4	10.0		41	5.0	5.1	5.0	7.7
B	7	10.0	10.0	10.0	10.0		42	5.0	6.5	6.3	6.2
	8	0.0	0.4	0.0	0.0	J	43	5.0	0.4	3.0	2.8
	9	5.0	3.1	4.7	3.9		44	5.0	7.5	4.6	5.2
	10	0.0	0.9	0.0	0.0		45	8.8	3.2	9.3	8.2
C	11	10.0	8.4	6.0	8.0		46	8.8	−8.4	6.5	6.8
	12	10.0	10.6	10.0	10.0	K	47	8.8	−2.1	6.5	6.8
	13	0.0	0.1	0.0	0.0		48	10.0	4.7	8.9	7.8
	14	10.0	8.1	8.2	10.0		49	0.0	−0.2	0.0	0.4
D	15	5.0	6.1	5.0	7.6		50	0.0	−0.2	2.5	0.8
	16	10.0	8.9	7.5	9.1		51	10.0	0.7	7.5	9.0
	17	5.0	0.7	0.0	0.0	L	52	6.9	−2.0	5.9	3.4
	18	5.0	2.7	1.5	3.4		53	8.1	−4.0	6.7	5.4
E	19	5.0	6.7	3.2	6.9		54	5.6	2.9	4.5	4.8
	20	5.0	8.9	5.0	10.0		55	2.5	−2.1	2.3	2.1
	21	10.0	7.5	2.5	6.9		56	10.0	7.3	7.9	7.5
	22	3.8	5.5	7.5	3.5	M	57	5.0	−2.0	10.0	9.2
F	23	2.5	3.1	4.3	3.9		58	0.0	−0.2	0.0	0.4
	24	0.0	0.3	1.7	0.0		59	10.0	4.8	8.4	8.5
	25	6.9	8.4	2.5	10.0		60	6.2	0.0	5.1	6.2
	26	10.0	10.2	10.0	10.0		61	2.5	7.7	2.9	6.6
	27	0.0	1.0	0.0	0.0	N	62	0.0	−0.1	2.5	3.6
G	28	10.0	9.4	10.0	7.3		63	10.0	0.7	7.5	8.3
	29	10.0	10.2	10.0	9.1		64	1.2	1.7	2.7	3.2
	30	10.0	9.4	10.0	8.9		65	5.7	0.5	5.0	5.5
	31	0.0	1.4	2.5	0.0		66	0.0	0.3	1.0	0.9
	32	7.5	6.6	6.0	8.4	O	67	10.0	4.8	10.0	9.2
H	33	5.0	6.0	4.4	6.6		68	1.4	2.6	1.4	2.0
	34	2.5	3.8	3.1	4.3		69	10.0	0.2	10.0	8.3
	35	10.0	9.5	7.5	10.0						

between these two extremes. The means were standardized to a 0–10 scale within each family, and appear in Table I, under Test 1.

This graphic rating-scale was used because it is rapid, it is easy to explain, and it is easy to score. The end-anchoring serves two purposes: first, it forces a uniform range of judgment on all *O*s; and second, it acts to overcome the strong 'bunching' tendency which frequently characterizes rating-scales. The procedure raises some questions, however, about the analysis of the resulting scales.

(1) Correlation coefficients computed from such scales (and manipulations of such coefficients, as in factor analysis) must be treated with caution, since the scattergrams are not bivariately normal; precautions were employed and are discussed where appropriate (*vide infra*).

(2) The direct use of means, instead of medians, or any form of the 'law of comparative judgment,' must be justified as a scaling procedure.[14] This was done in three ways: by comparison with other (validating) measurement-procedures; by linear plotting against the values given by the predictive equation (cross-validation); and by the cross-family comparisons, which should be quite sensitive to possible distortions of scale, since any distortions within families should tend to scramble the relative order of 'inner' members in cross-family comparisons, when several sets of such members are combined in one series.

(3) The end-anchoring could cause misleadingly high correlations in the following manner: in each family, only two figures might have been chosen which were clearly more and clearly less three dimensional, respectively, than the remaining figures; even if this difference itself was small, the *O*s were forced to place the two figures at the extremes of the scale. This possibility was rejected by testing the over-all correlation after removing the scores for the extreme pair of figures in each family.[15]

Results. Seventeen tests were devised on which to scale the 40 stimulus-figures; they are numbered 2–18 (the *O*s' transformed mean responses being counted as Test 1). The tests were chosen for geometric variety and for relative ease and reliability of scoring—although some of them still seem almost too laborious for practical use, especially Tests 14 and 16, which had been chosen as simpler substitutes for such preliminary measures of *symmetry* as 'average difference between cross compared angles within each subfigure.' The tests follow:

Test 2: Total number of interior angles.
Test 3: Total number of line segments (*i.e.* counting a new segment at each inflection-point, intersection, and so forth).
Test 4: Total number of continuous line-segments (*i.e.* ignoring intersections).
Test 5: Total number of points of intersection of two or more lines.
Test 6: Total number of points of intersection of three or more lines.
Test 7: Total number of outside (perimeter) line-segments.
Test 8: Total number of inside line-segments.
Test 9: Test 8 divided by Test 7.
Test 10: Test 8 divided by the sum of Tests 7 and 8.
Test 11: Total number of different angles (differing by at least 1°); this was determined from the master patterns from which the experimental stimuli were reproduced, because of the difficulty of calculating it geometrically for the more complex figures, and because of the possibility of errors in construction.
Test 12: Test 11 divided by Test 2.

---

[14] J. P. Guilford, *Psychometric Methods*, 2nd ed., 1954, 37, 223-244, 288 f.
[15] This method was suggested by Professor M. E. Bitterman.

Test 13: Total number of degrees in all interior angles (*i.e.* sum of angles to nearest degree).

Test 14: Sum of the differences, in degrees, between the measured value of each angle, and what that angle would have been if the sub-figure had been equi-angular (*e.g.* for a 30°-60°-90° triangle, this score would be 30° — 60° = —30°, plus 60° —60°, plus 90° — 60° = 30°, totals 60°).

Test 15: Test 14 divided by Test 2.

Test 16: The sum of the absolute differences, in degrees, between 'adjacent angles' (*i.e.* angles which fall to either side of a line intersecting another line or lines).

Test 17: Total number of different closed sub-figures in each pattern.

Test 18: Test 16 divided by Test 2.

The scores on each test (including the scores on Test 1, which is the scale of the means of the *Os*' responses to each figure) were transformed to

TABLE II

CORRELATION-MATRIX OF TESTS 1 (APPARENT TRIDIMENSIONALITY) AND TESTS 2–18 (MEASURES OF BIDIMENSIONAL COMPLEXITY)
(Decimal points omitted)

Test	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
2	61																
3	71	89															
4	89	53	67														
5	71	93	86	60													
6	65	92	91	60	88												
7	51	47	59	40	59	50											
8	61	99	89	53	93	92	47										
9	42	83	73	44	71	74	13	84									
10	40	80	72	45	68	71	15	80	87								
11	63	61	48	52	59	47	03	62	59	56							
12	48	16	17	42	24	11	−20	18	23	23	73						
13	53	90	76	50	86	77	39	90	79	75	54	09					
14	45	55	52	53	52	43	07	55	73	69	59	32	62				
15	54	20	27	63	25	16	−09	21	37	35	61	67	20	69			
16	52	78	64	41	70	66	12	78	76	69	82	72	67	59	30		
17	40	66	56	43	56	64	20	64	65	60	35	06	64	59	28	58	
18	45	65	53	35	57	55	−04	65	66	59	78	52	54	47	29	94	48
Test	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

a 10-point scale, the highest-scoring figure *within each family* being 10, the lowest being 0. The scores in Test 1 were then correlated with each of the other sets of Test-Scores. The coefficients of correlation are given in the first column of Table II. All of these coefficients are positive, varying from 0.40 to 0.89. Since the relationship hypothesized is a simple function, the product-moment coefficient of correlation was employed to measure all dependencies. Because of the peculiar nature of the bivariate distributions, all significance tests were, however, made with rank-difference correlations.

Discussion. The relationship hypothesized appears to hold, as a first approximation, for all of the dimensions which were investigated here: the

greater the two-dimensional complexity, the greater the apparent three-dimensionality of the figure. If a single stimulus-dimension had to be chosen at this point to predict *Os'* responses, it would be *Test 4,* and the predictive equation would be:

$$Y = K\,(T_4) + C, \ldots\ldots\ldots\ldots \text{Equation [1]}$$

where Y is the response to each individual figure (within a family) on a 0–10 scale (Test 1), T_4 is the score of that figure on Test 4 (number of continuous line-segments), K is very close to being a constant of value 1.0 in the ranges investigated here, and C is a constant of low value, approximately zero. Equation [1] predicts the scores of Figs. 1–40 with $r = 0.89$, and predicts the scores of Figs. 1–20, with $r = 0.85$.

Does some 'higher-order,' psychophysical variable cut across Tests 2–18, *i.e.* can we improve on Test 4? Elucidation of such an hypothetical predictive

TABLE III

Correlations of Equations [1], [2], and [3] with Test 1 (Obtained Mean Scaled Responses)

Equation	r_{1-20}	r_{1-40}	r_{41-60}
[1]	0.85	0.89	0.80
[2]	0.98	0.96	0.09
[3]	0.91	0.80	0.98

variable, or set of variables, depends upon reduction of Table II, which was attempted by: (*A*) multiple-regression analysis of Table II; (*B*) factor analysis of Table II; (*C*) cross-validation of the Equation [1] above, and of the equations resulting from the multiple-regression analysis, and from the factor analysis.

As noted earlier, the coefficients of Table II do not arise from bivariately normal distributions; moreover, the matrix of tests which they represent contains many pairs which are not independent, by the nature of their calculation (for example, Tests 11 and 12 must yield some index-correlation). For this reason, and because of the necessarily-arbitrary nature of the sample of stimuli upon which all of these tests were based, the weights found by multiple-regression analysis (Equation [2], below), and the factors displayed in the body of Table IV must be considered solely as exploratory aids, rather than stable entities. No interpretation of either of these sets of numbers can be attempted without their cross-validation, employing a new set of stimulus-figures, which was undertaken in Experiment II). We should also note here, however, that the factors finally drawn from Table IV (specifically, those with heavy loadings in Test 1) cannot be

artifacts caused by the lack of independence of many of Tests 2–18, for reasons given below.

(A) Multiple-regression analysis. A multiple-R of 0.96 was yielded by the following multiple-regression equation:

$$Y = -0.167 + 0.85X_2 + 0.32X_3 + 0.53X_4 + 0.39X_5 + 0.01X_6 + 0.11X_7 - 0.47X_8 + 0.09X_9 - 0.09X_{10} + 0.60X_{11} - 0.00X_{12} + 0.01X_{13} - 0.17X_{14} - 0.06X_{15} - 1.19X_{16} + 0.12X_{17} - 0.05X_{18} \quad \text{...Equation [2]}$$

Table I gives the values 'predicted' by this equation, compared to the obtained responses; Figs. 1–20 are the projections of 'single objects' (*i.e.* those composed of connected plane surfaces, in their three-dimensional versions). For the entire set of figures, Figs. 1–40, $r = 0.96$; for the subset, Figs. 1–20, $r = 0.98$.

TABLE IV
FACTOR ANALYSIS OF TESTS 1–18, FIGS. 1–40
Rotated factor loadings (decimal points omitted)

Test \ Factor	1	2	3	4	5	6	7	8
1	620	−292	691	111	−019	−021	035	000
2	999	016	011	−105	−067	−068	−011	−043
3	905	078	275	−012	012	294	−025	001
4	566	−196	618	337	007	029	166	−050
5	929	−029	236	−077	147	002	−193	043
6	915	094	189	−128	043	113	160	067
7	418	335	534	−198	034	−001	−316	011
8	996	002	001	−113	110	−069	−006	−026
9	873	−058	−270	252	−036	056	062	−029
10	837	−039	−238	241	038	094	052	−098
11	615	−749	021	093	−038	−161	−114	−016
12	198	−850	118	171	075	103	081	006
13	905	109	−019	080	038	−239	−076	007
14	628	−159	−063	649	−091	−017	−095	−018
15	263	−524	224	678	094	003	021	034
16	812	−429	−152	−087	−395	000	−078	−002
17	687	103	−025	225	−192	−010	164	040
18	683	−518	−180	−124	−338	016	117	−014

Three disadvantages apply, despite the high coefficients of correlation, to this multiple-regression equation: (1) Inspection of the weights, in Equation [2], does not suggest any rational pattern (probably because of the considerable amount of index-correlation involved); since the factor analysis, described below, was immediately more informative, and because of the other disadvantages of this regression-equation, no attempt was made to simplify or reduce the number of its terms.

(2) The underlying principles expressed by this equation are unclear, since the weights are specific to the set of tests employed; moreover, one would assume that almost any set of 17 arrays of scores probably could be suitably combined to pre-

dict equally well the particular set of 40 scores obtained. In support of this assertion, we may note that this equation failed to predict the response-scores to a new set of stimulus-figures in the cross-validation procedures of Experiment II (see Plate II, Figs. 41–69).

(3) The weights in the regression-equation are also likely to be specific to the scoring procedure employed, and it would be desirable to have an equation which could be applied to the results of any other scoring procedure. Scoring for the equation is laborious and expensive, and this laboriousness makes it very difficult to extend further the ecological sample.[16] These disadvantages prompted the performance of the factor analysis.

(B) Factor analysis. Table II yielded the factor-loadings (rotated, Quartimax, IBM 650) in Table IV. The first three factors appear to account for 94% of the variance of our psychophysical response-scores. For Test 1—mean tridimensionality of each figure within each family—Factors 1, 2, and 3 yield $h^2 = 0.94$. For these factors, the highest loadings for each, in Tests 2–18 are: (*Factor 1*) Tests 2, 8, 5, and 6, all positive (tentative names—number of angles, intersections or angular complexity, crossings); (*Factor 2*) Tests 12, 11, 15, and 18, all negative, as is loading for this factor in Test 1 (tentative names—mean or specific angular asymmetry or diversity); (*Factor 3*) Tests 4, 7, 3, and 9 (tentative names—linear complexity or discontinuity). These factors seem quite easy to identify: (1) angular complexity; (2) specific angular asymmetry or diversity (*i.e.* number of different angles divided by number of angles); and (3) number of continuous lines (or linear discontinuity).

Since Tests 2–18 contain only general and specific factors and no error variates (*i.e.* these 'tests' are 'completely' reliable), since index-correlation can occur only between Tests 2–18, and since we only consider those factors which appear with high loadings in Test 1 (the 'criterion'), it seems safe to assume that Factors 1–3 are not spurious artifacts arising through our use of dependent measures (*i.e.* the index-correlation of tests such as 11 and 12).[17] Moreover, since $h^2 = 0.94$ for the first three factor-loadings in Test 1, neither spurious factors due to index-correlation, nor non-linearity due to this procedure or other causes, would seem to be responsible for the loss or suppression of factors predictive of the criterion.

Now we wish to obtain a predictive equation from Table IV. The tests most highly loaded in these three factors are, respectively, Tests 2, 12, and 4. Test 4 appears to be closest to factor loadings in Test 1, the 'criterion,'

[16] Brunswik, *op. cit.*, 26-55.

[17] A. O. H. Roberts, "Artifactor"-analysis: some theoretical background and practical demonstrations, *J. National Institute for Personnel Research* (Johannesburg), 7, 1959, 168-188; J. S. Moritz, Note on "Artifactor"-analysis, *ibid.*, 7, 1959, 191.

and is more highly predictive of Test 1 than either Test 2 or Test 12 (see Table II). Accordingly (and arbitrarily), we shall weight it more heavily. Since there are some families of figures for which any given test will yield the same score for all projections (*i.e.* will fail to differentiate—for example, there are the same number of continuous line-segments in all of the figures in Family *J*, Plate II), we must so set up our equation that the relative apparent tridimensionality predicted will in such cases be determined solely by the remaining tests in the equation. Because it would be convenient to keep the equation in an additive form (which allows other terms to be added when testing new types of material), we shall simply divide the sum of the weighted test-scores by the number of tests which are actually contributing to the discrimination of the members of any given family. In symbols, this gives us:

$$Y_i = [\Sigma(T_2 + T_{12} + 2 \times [T_4])_i]/N_T, \ldots\ldots \textit{Equation} \quad [3]$$

where Y_i is the predicted apparent tridimensionality-score for any figure (on a scale of 0–10); T_2, T_{12}, T_4 are scores on the appropriate tests, for that figure relative to the rest of its family; and N_T is the number of tests (with $2 \times T_4$ counted as $N = 2$) which actually discriminate between the figures of the family being considered. The values yielded by substituting the measurements of each figure in Equation [3] are given in Table III.

Conclusions. The sizable correlations between the values yielded by each of these equations (and the correlations given in Table II), support the first hypothesis, *i.e.* that the relative apparent tridimensionality of each member of a family of reversible-perspective representations of a given three-dimensional object will be a simple function of the geometrical complexity of the two-dimensional stimulus (as a first approximation). Three specific equations for relating apparent tridimensionality to particular measures of complexity were selected from the present data, with correlations ranging from 0.85–0.98 for Figs. 1–20, and from 0.80–0.96 for Figs. 1–40. The simplest equation, Equation [1], was the single test most highly correlated with apparent tridimensionality. The multiple-regression equation, Equation [2], involving all of the 17 measures of geometric complexity, yielded scores most highly correlated with apparent tridimensionality, but the use of this equation entailed several disadvantages. Equation [3], involving three of the measures of complexity, was based upon a factor analysis of the correlation matrix of Tests 1–18; it is both rational and relatively simple to use. It appears to be the equation to adopt (although all three equations are more-or-less well fitted to the response data). We still

do not know, however, whether this equation will predict responses to a totally new group of stimulus-figures of the same class. Accordingly, Experiment II was designed to provide *cross-validation.*

Experiment II. Cross-Validation With New Stimulus-Figures

Problem. To determine the generality of the equations developed in Experiment I, in predicting the relative apparent tridimensionality of the class

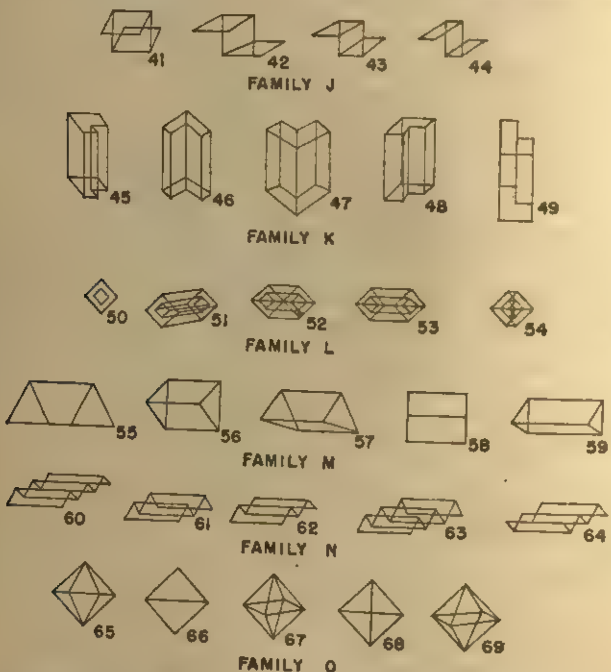


Plate II. Stimulus-Figures Used for Cross-Validation

of stimulus-figures represented by Figs. 1–20 (*i.e.* the reversible-perspective projections of single objects composed of connected plane surfaces).

Method. The *Os*, of the same categories as in Experiment I, were divided into two groups. Group A, 25 *Os*, were tested individually; Group B, 150 *Os*, were tested in a group-administered experiment. None of the *Os* in Experiment II had participated in Experiment I.

For Group A, a new set of stimulus-figures was devised. The 29 stimulus-figures of Families J–O (Plate II) were chosen to meet the criteria for selecting Figs. 1–20 in Plate I. Procedures were as in Experiment I, except that the test-pages were not stapled into booklets and were individually administered.

The *Os* of Group B arranged small, rectangular cut-outs of the stimulus-figures

along a one-foot scale running from 0–10, and then recorded the results. The specific stimulus-figures used were selected from Families A, B, K, and O; they were Figs. 1–4 and 5′–8′ (the prime signs indicating that the major axes of these figures were changed to the vertical) from Plate I, and Figs. 45, 47, 48, 49, 65, 66, 67, and 68 from Plate II.

The instructions for both groups were the same as those employed in Experiment I.

Results: (1) Group A. In Experiment I, it will be recalled, all response-scores were normalized to a 0–10 scale for each family in order to weight all families equally in devising the predictive equations. In the present experiment, we used the simple unadjusted means of the ratings of each figure. These means are given in Table I. Their correlation with the predicted values from Equation [3] is $r = 0.98$. To test these predictions, we may note that the rank-order correlation coefficient, $r_s = 0.98$, $t = 13.0$, $p < 0.001$. Equation [1] (Test 4, alone) predicts the responses to these new figures with an r of 0.80. Contrasted with the ability of Equation [3] —based on the factor analysis—to predict responses to new reversible-perspective forms, Equation [2]—the multiple-regression equation—falls to a value of $r = 0.09$, with these data.

(2) Group B. In Group B, Figs. 1–8 comprise a modified retest of Families A and B in Experiment I; Figs. 2 and 4 have smaller central quadrilaterals than those of Experiment I, and Figs. 5′–8′ have been rotated 90° from their orientations in Experiment I. Figs. 45, 47, 48, 49, 65, 66, 67, and 68 provide a retest for the corresponding stimulus-figures of Group A. In all cases, a different scaling procedure (see above) was used than that employed in Group A. For these 16 stimulus-figures, the obtained means were 9.5, 6.8, 9.1, 0.0, 4.9, 9.4, 10.0, 0.0, 9.4, 0.0, 10.0, 5.8, 4.9, 1.7, 10.0, and 0.7, respectively; their respective predicted values may be found in Table I.

Discussion. Of the three equations fitted to the data in Experiment I, Equation [3]—based upon the factor-analysis—displays for the *O*s of Group A, the best over-all performance in cross-validation, ($r = 0.98$), *i.e.* in predicting the apparent tridimensionality of a new group of stimulus-figures. Equation [1]—Test 4—is next best ($r = 0.80$), while Equation [2]—the multiple-regression equation which employs the weighted scores on all 17 tests—fails to predict the responses to the new figures ($r = .09$).

In determining the scale of apparent tridimensionality with the *O*s of Group A, mean scale-positions were employed, without rescaling the means to a 0–10 range. Such normalization was a necessary constraint in the exploratory procedures of Experiment I. To retain this constraint on varia-

bility in the present experiment would have been to run the risk of having achieved a psychophysical equation which would be too narrow and which would be generalizable only to within-family relationships; the unconstrained mean judgments might have ranged from 3 to 6 in one family, and from 0 to 10 in another, yet we would have had no way of knowing it. It so happens that—as our present results show—this is not the case.

Each individual *O* was still constrained, however, to end-anchor his *judgments* at 0 and 10, in order to overcome individual differences in readiness to use the entire range of each scale. As discussed in connection with Experiment I, this might lead to the appearance of a general relationship between stimulus- and response-scales, even if it were only the extremes within each family which were predictable from Equation [3]. This possibility was examined by discarding the two extreme stimuli within each family, and measuring the correlation between the remaining data from Experiments I and II (Families A–E, J–O) on the one hand, and the values obtained from Equation [3] on the other. The correlation is 0.81, with $r_s = 0.80$, $t = 6.6$ and $p < 0.001$. This comparison is essentially the measurement of a cross-family relationship, and we will return to it later when we consider such relationships.

At present, let us note that the cross-validation procedure has supported the utility of Equation [3]. We now have evidence of the inter-stimulus and the inter-*O* reliability of the equation, but no measure of test-retest reliability, nor of the consistency of such findings under different conditions of judgment. It was for these purposes that Experiments III and IV were performed.

Experiment III: Test-Retest Reliability

Families A and E (chosen because they comprised the most reliable and least reliable of the families among Figs. 1–20) were re-administered to 50 of the original *O*s after two months' elapsed time, interspersed in balanced orders between other families of figures. The correlation between the mean ratings for first and second presentations for these eight figures was 0.99.

Experiment IV: Generality of Scaling Procedures

Would the same simple pattern of results have been obtained if different measurement-procedures had been employed, or is the efficacy of the predictive equation restricted to the judgmental procedures under which it was derived?

Procedure. Thirty new *Os*, students of introductory psychology and naïve as to the purposes of the experiment, repeated the procedures of Group B, in Experiment II, except that they performed a paired-comparison of labels (numbers) of each stimulus-member of each of the four families, instead of using the end-anchored graphic rating-scale.

Results. Rank-orders of apparent tridimensionality were obtained from the paired-comparisons data, and are shown in Row 1, Table VA (there was too much agreement between judges to attempt any interval-scaling by use of Thurstone's methods[18]). The mean ratings from Group B, Experiment II, were converted to rank order within families, and are shown in Row 2. Of the 16 pairs of ranks, there is only one single-place reversal.

Discussion. Two additional sets of data are relevant to this question, having been gathered several years ago by Levitt and Hochberg in the course of an unpublished study and by McAlister in the course of his doctoral research.[19]

In the study by Levitt and Hochberg, 10 of the 11 stimulus-figures of Families A, B, and D were shown within one single group of shapes. Each of 20 *Os* pressed a telegraph key when the figure appeared tridimensional, and released it when the figure appeared flat, for a total of 10 reversals. They also ranked the figures from most two-dimensional to most three-dimensional (the two tasks were presented in balanced order). In McAlister's study, 17 of the 20 forms used here in the first five families were presented amidst a total of 39 figures; *Os* judged the entire series of figures by two methods, the method of successive categories and a rank-rate method, similar to the one employed in Experiments I and II.

These two sets of data are of special interest to us here because judgments were obtained for each figure in comparison with an entire series of figures from different families, and it might be that the within-family comparison we have employed here generated a spurious set of relationships, or perhaps the within-family comparison tapped real relationships which are too labile or too fragile to withstand the interactions and anchoring effects which would result from embedding the figures in a full series, out of 'familial context.' A comparison of our present data with those of these two studies, will answer this question. The differences in the methods of measurement used in these experiments and the present ones make it necessary, however, to perform some sort of transformation of the earlier data: our psychophysical equations at present are equipped only to generate comparisons

[18] L. L. Thurstone, A law of comparative judgment, *Psychol. Rev.*, 34, 1927, 273-286.
[19] McAlister, *op. cit.*, 37-72.

within families. If we temporarily ignore the fact that an entire series of shapes was treated as a unit in both of these previous series of judgments (instead of being judged only within families as we have done in the present experiments) we can then convert the position of each stimulus-figure to a scale running from 0–10 within each family, and compare the previous data to the present findings.

In Table VB, the various methods are compared as to the fit of their data to those of the present experiment. The correlation of the present results with the telegraph-key time-scores is 0.99, and with McAlister's data, 0.77.

Despite these high correlations, however, we cannot claim the ability to predict cross-family judgments. We should note the reasons why our present

TABLE V

Results of Experiment IV and Supplementary Studies

(A) Experiment IV

Ranked by	Stimulus-figures															
	1	2	3	4	5′	6′	7′	8′	45	47	48	49	65	66	67	68
Method of paired comparison	4	2	3	1	2	4	3	1	3	1	4	2	3	2	4	1
End-anchored scale (Experiment II)	4	2	3	1	2	3	4	1	3	1	4	2	3	2	4	1

(B) Supplementary studies

Ranked by	Stimulus-figures									
	1	2	3	4	5	6	7	13	14	16
Ranks (reranked within families)	1	3	2	4	1	2	3	1	3	2
Present data (Exp. I) ranked	1	3	2	4	1	2.5	2.5	1	3	2
Telegraph-key time-scores, normalized	0.0	5.8	5.3	10.0	0.0	10.0	10.0	0.0	10.0	8.6
Present-data, re-normalized within families	0.0	7.5	5.3	10.0	0.0	10.0	10.0	0.0	10.0	9.2

psychophysical equation makes no provision for cross-family comparisons: its terms include only measures of two-dimensional stimulus-complexity, since it was devised only for and from the comparison between stimulus-figures having identical tridimensional complexity. Research employing cross-family comparison procedures will be necessary before we can attempt to formulate a new equation which is applicable to groups of figures which vary both in two-dimensional and tri-dimensional complexity. Nevertheless, we can see at this time that all within-family predictive comparisons, regardless of the family being considered, are being made on the same approximately-linear scale, and we have just seen that very similar scales are obtained in cross-family judgments. Consequently, prospects seem good for achieving a more flexible predictive equation, applicable to such general cross-family comparisons.

CONCLUSIONS

We now see that even ambiguous figures, selected for their reversibility-of-depth, can be submitted to psychophysical investigation, and can yield (a) high intra- and inter-*O* reliabilities (Experiments II and III), (b) lawful relationship to stimulus-characteristics (Experiments I and II), and (c) generalizability across different methods of measurement (Experiment IV). Consequently, the psychophysical data may be summarized by empirical and rational 'rules' or equations (Experiment I) which will predict the responses obtained to future samples of stimulus-figures (Experiment II).

Specifically, we can prescribe a basis for choice between a single object's projections, when considering alternative reversible-perspective representations: the average person's judgment of the relative solidity of alternative (reversible-perspective) versions of a given single object may be predicted in advance from Equation [3]. The present recommendation is that (other things being equal) such figures should be constructed with as many angles, as many different angles, and as many different continuous lines as will fit the limits set by the projection-family of the represented tridimensional object.

We should note that we have demonstrated that one can find useful psychophysical relationships, even in traditionally 'indeterminate' stimulus-response areas, without invoking either unmeasured cortical processes of 'organization,' or equally unmeasured 'past experiences.' Moreover, the findings are not 'merely empirical.' The subtleties of the Gestalt relationships need not be irretrievably lost through the processes of statistical quantitative analysis; if the Gestalt factors are operative in determining the psychophysical relationships at all, they must reappear in some form in the results of analysis. In the present investigations, the factors which emerge are readily identifiable as approximations to the more intuitive 'laws' which have been reported in Gestalt writings:

Factor 1: Number of angles. This is approximately equivalent to 'complexity,' or in its inverse, 'simplicity.'

Factor 2: Number of continuous line-segments. For any given object whose various projections are being considered, the smaller the score in this factor, the better the 'continuation'; therefore, the inverse of this is the traditional 'good continuation.'

Factor 3: Number of different angles, divided by the total number of angles. Inspection of Tables II and IV suggests that this factor largely measures the inverse of 'symmetry,' *i.e.* the lower this score for any of the various projections of a given object, the greater the symmetry of that projection.

In short, the three-factor equation predicts the relative three-dimensionality of the various members of a family of reversible-perspective projections of a given object in the following way: the greater the *complexity*, the *asymmetry*, and the *discontinuity* of the projection of a given tridimensional object in two-dimensions, the more three-dimensional it will appear. We may, in reality, be dealing with only one dimension—'figural goodness'[20] which appears here as three independent factors only because our measuring procedures cut different aspects of it, but such speculation is premature at present, since the 'ecological generality' (the extent to which this equation can be used to predict responses to *other* classes of representational figures) is as yet unknown. Most probably we do not yet have the 'general law of line drawing,' much less one which will subsume all of perception, as any attempt to force unitary laws at this time would imply.

There are two clear directions in which to proceed. The first is to attempt to raise the predictive power of our psychophysical equations as far as error-of-measurement and subject-variability will permit. The second is to seek to extend the variety of stimulus-situations for which prediction can be made, to widen further the range of stimulus-samples, and to determine well-defined categories of stimulus-situation for which different classes of psychophysical equations will predict with differential success. Of these two alternatives, the latter appears to be the sounder strategy: only when such categories include the majority of stimulus-situations with which men must deal, from graphs and maps to landscapes and portraits, will it be appropriate to seek unifying 'principles of perception,' and to seek 'perfect' psychophysical equations.

Summary

Factor-analytic techniques assisted in obtaining a psychophysical equation which will predict the within-family responses which untrained observers make to reversible-perspective projection drawings of three-dimensional objects. Populational and procedural generality of the results were measured, and they were found to be in good agreement (*r*s between 0.7 to 0.9). Practical and theoretical implications are discussed.

[20] Koffka, *op. cit.*, 151-176.

RESPONSE-FACTORS IN LEARNING AND TRANSFER

By Bennet B. Murdock, Jr., University of Vermont

In a previous experiment it was found that the type of response used in a paired-associate task had a significant effect on both learning and transfer.[1] Verbal responses (100% association-value nonsense-syllables) were easier to learn than motor responses (pressing the correct push-buttons), all other factors held constant.[2] Also, even though the two types of response had been learned to the same criterion, verbal response led to significantly more transfer in a second task whether the responses themselves in Task 2 were verbal or motor. To account for these findings it was suggested that verbal cues produced by the responses were more distinctive than motor cues, and that the difference in the distinctiveness of the cues resulted in the differences in both learning and transfer.

This paper reports the results of six experiments which were conducted to obtain further information on response as a factor in learning and transfer. Experiment I was a reaction-time (*RT*) study comparing verbal and motor responses in choice *RT*. Experiments II and III were essentially replications of the original study using a somewhat different method. Experiment IV was a test of the assumption that verbal and motor responses differ in the distinctiveness of their response-produced cues. Experiment V tried to determine if responses of known differences in distinctiveness would result in differences in transfer. Experiment VI was a direct study of mediating responses.

Experiment I

This experiment, on choice *RT*, was conducted to determine if the differences in learning previously reported could be attributed to differences in speed of reaction. The paired-associate task, in which the verbal responses were learned more rapidly than the motor responses, had used a rather

* Received for publication September 22, 1958. This work was supported by a research grant, G-2590, from the National Science Foundation. Mr. James Barnard tested *Ss* and helped in the analysis of the data.

[1] B. B. Murdock, Jr., Effects of task difficulty, stimulus-similarity, and type of response on stimulus predifferentiation, *J. exp. Psychol.*, 55, 1958, 167-172.

[2] The terms 'verbal' and 'motor' are used throughout for the sake of convenience to describe the two types of responses, and do not imply that the specific responses studied are an adequate sample of some larger population of responses that might bear these labels.

rapid rate of presentation (1.5 sec. for the stimulus-object (S) and 1.5 sec. for the stimulus-response S—R). The group learning the motor reactions might have been at a disadvantage merely because they were not able to execute the motor responses in the limited time available.

Procedure. The stimulus-objects in this choice RT experiment were eight lights (2-w. neon glow lamps) spaced horizontally 2 in. from center to center. They were located 14 in. above a table that was between the subject (S) and the experimenter (E), and were approximately 24 in. from S. The eight keys for the motor reactions were telegraph keys located on the table before S, and were arranged in two semicircles of four each so S could comfortably place the four fingers of each hand on the keys (the thumb was not used).

On each trial, following a 'Ready' signal, one of the eight lights came on. On the motor trials S depressed the correct key as quickly as possible. On the verbal trials S called out the number of the light, speaking into a microphone he held in his hand. The S was instructed that he was to consider the lights and the keys as being numbered from 1 to 8 (left to right). The correct motor response to any light was to press the key of the same number. The correct verbal response was to say the same number as the light. Thus, when Light 5 came on S pressed Key 5 on a motor trial or said 'five' on a verbal trial.

Following 16 practice-trials (in which each reaction, 8 verbal and 8 motor, occurred once) each S was given 96 test-trials. For all Ss the order was VMMVVMMV . . . V throughout the entire 96 trials. This was done to control for any practice- or fatigue-effects. With the 96 trials each of the 16 reactions occurred 6 times, and the order of presentation was determined from a table of random numbers. Four different orders of presentation were used. Every 24 trials S was given a 2-min. rest.

There were 16 Ss, 8 men and 8 women, obtained from the first-year course in psychology.

Results. The main results are shown in Fig. 1, where the mean RT for the verbal and the motor responses to each of the eight lights are plotted. Each S's reaction to each light was the mean of six trials, and each point in the graph is the mean value for 16 Ss. For all of the lights the mean verbal RT was slower than the mean motor RT.

To test significance, for each S the verbal and motor responses were paired. The first verbal and first motor response to Light 1 were paired, the second verbal and second motor response to Light 1 were paired, and so on up to the sixth verbal and sixth motor response to Light 8. Thus, for each S there were altogether 48 pairs. For simplicity the sign-test was used, and for every one of the 16 Ss there were significantly ($P < 0.05$) more cases where the verbal RT was slower than where the motor RT was slower. To generalize to the population of Ss, for each S a mean difference (verbal-motor) was obtained. The mean of the 16 Ss was +139 m.sec., and this was significantly different from zero ($t = 8.32$, $P < 0.01$). There were no

significant sex differences in any of the results. Thus, each *S* individually as well as all 16 *Ss* taken as a group showed a longer *RT* when a verbal response was required.

These findings agree with the results of a study by Alluisi and Muller.[3] They found motor responses to be faster (though less accurate) than verbal responses in a self-paced serial task, irrespective of which of seven visual symbolic codes was used. It seems most unlikely, then, that the faster learning of verbal responses could be attributed to a longer period of time required to execute the motor responses.

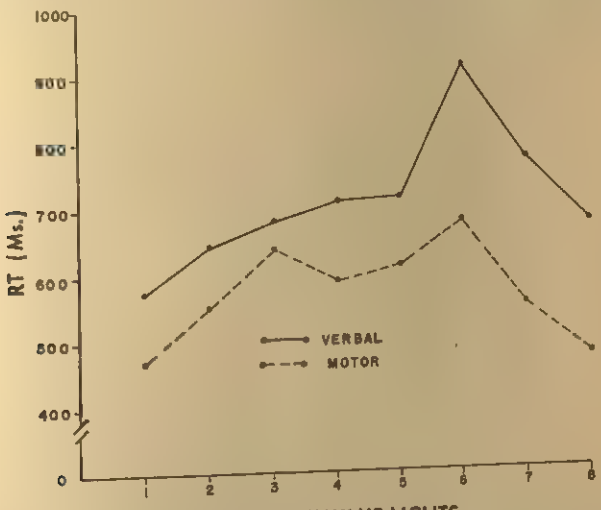


FIG. 1. MEAN REACTION-TIME TO EVERY STIMULUS-CONDITION IN EXPERIMENT I

EXPERIMENT II

The purpose of Experiment II was to replicate a part of the previous study, in which motor responses were found to be harder than verbal responses, and to determine if, under somewhat different conditions, the same results would occur.[4] The verbal responses used in this experiment were numbers rather than nonsense syllables. There were eight stimulus-numbers and eight responses, so in the verbal group *S* learned to say the numbers

[3] E. A. Alluisi and P. F. Muller, Jr., Verbal and motor responses to seven symbolic visual codes: A study in S—R compatibility, *J. exp. Psychol.*, 55, 1958, 247-254.

[4] Murdock, *op. cit.*, 170.

one through eight to the appropriate numbers while in the motor group *S* learned to press Keys 1 through 8 to the same numbers. Thus, the only difference between the two conditions was in the way the response was executed; given a particular number, *S* either had to say "three" or press down Key 3.

Procedure. To give the results greater generality, many of the conditions of the previous experiment were deliberately changed.[5] The stimulus-objects in the previous experiment were patterns of lights in a 5×5 matrix. Those employed here were patterns of lights in a 6×2 matrix; the particular patterns used are shown

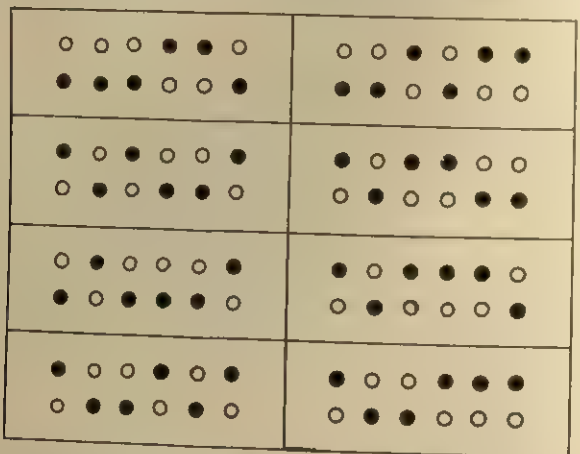


FIG. 2. STIMULUS-CONDITIONS IN EXPERIMENT II
A filled circle represents a light on, an unfilled circle a light off.

in Fig. 2. With this matrix and the rule that one light in each column can be used, there are a total of 2^6 possible patterns. The sample of 8 was selected in the following way from the population of 64. Some patterns were discarded as being too easy. Those discarded were: all patterns with five or six lights on in either row, all patterns with four consecutive lights on in either row, and all patterns with left and right halves either identical or mirror images. Of 64 patterns this left 32; these 32 were divided into four groups of 8 each and were used as stimulus-patterns in a pilot experiment with 8 *Ss*. The results indicated that these four groups were comparable in ease of learning. The one of the four groups selected is shown in Fig. 2.

The task used a self-paced, forced-anticipation method. A given stimulus-pattern was presented on the first trial and *S* could take as long as he wished to respond. Then, after this response a signal light came on for 2 sec. which indicated what was the correct response for that stimulus. On subsequent trials the *S* had to

[5] On the rationale of this procedure see B. J. Underwood, *Psychological Research*, 1957, 281-287.

respond before the confirming signal light came on. The indicators of the correct verbal response were in a column directly above the stimulus-panel, and the number was written next to each signal light. The lights indicating the correct motor response were mounted on the same panel as the telegraph keys and each light was in line with its key. The signal lights and numbers of the verbal group were hidden for the motor group, and the response-panel was removed from the table for the verbal group.

All *Ss* learned to a criterion of two consecutive perfect trials, or worked for 45 min., whichever came first. The stimulus-patterns were programmed by *E*, who shuffled the cards indicating the order of presentation between each trial. Thus the order of presentation was random and different on each trial. The inter-trial

TABLE I

MEDIAN TRIALS TO SUCCESSIVE CRITERIA FOR *Ss* GIVING A MOTOR OR VERBAL RESPONSE TO IDENTICAL STIMULUS-CONDITIONS

Criterion	Motor	Verbal	Difference
1/8	1	1	0
2/8	2	1	1
3/8	3	2	1
4/8	5	3	2
5/8	8	5	3
6/8	10	7	3
7/8	16	11	5
8/8	23	14	9
2×8	29	15	14

interval was 10 sec. There were 30 *Ss* from the same population as before, and they were randomly assigned to the verbal or motor group.

Results. Of the 15 *Ss* in each group, 9 in the motor group but only 2 in the verbal group failed to reach criterion in the allotted time + 45 min. This difference was not due to the verbal group responding more rapidly, since the average time per trial was less for the motor group than for the verbal group (means of 70.2 and 80.4 sec. respectively, $t = 2.43$, $P < 0.05$).

To compare the groups, the unrealistic assumption was made that all *Ss* who had failed to reach criterion would have done so in the next two trials. This assumption underestimates the time required by the motor group for learning more than it does for the verbal group, and so works against the hypothesis under investigation. With this assumption, the median number of trials required to learn the verbal responses was 15, and the median number of trials required to learn the motor responses was 29. This difference was significant when tested by the Mann-Whitney *U* test ($U = 35$, $P < 0.01$).

The median number of trials to reach each successive criterion was determined in order to discover where in the learning period the greater diffi-

culty of the motor responses became manifest. These results are shown in Table I. It is apparent from the last column that the differences regularly increased with increasing degrees of mastery. Thus, the higher the criterion the greater the difference between the verbal and the motor group.

This experiment, then, confirms the previous study in finding motor responses more difficult to learn than verbal responses. This held despite the fact that, in a sense, the 'responses' were the same; the only difference was in how they were executed.

Experiment III

The purpose of this experiment was to replicate the previous study that found motor responses to result in less transfer than verbal responses. In

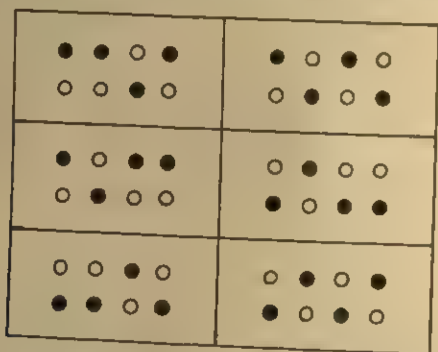


Fig. 3. Stimulus-Conditions in Experiment III
A filled circle represents a light on, an unfilled circle a light off.

Experiment III, as in Experiment II, many of the conditions differed from that of the previous study.

Procedure. The stimulus-patterns used in this experiment are shown in Fig. 3. There were only six stimulus-conditions and the matrix was 4 × 2 instead of 6 × 2. These changes were made in the hope that they would make the learning of the first task easier, so all *Ss* would be able to reach criterion. In the transfer task, Task 2, *S* had to learn into which channel to move the wobble stick of the Star-Discrimeter for each stimulus-pattern.[6] There were eight channels spaced every 45°; those in the 180° and 360° positions were not used.

The *Ss* in the verbal group learned the 100% association-value nonsense-syllables NOV, FAM, GUL, WIS, HOB, and PED as responses to the stimulus-lights, while *Ss* in the motor group learned to press Keys 1–6 to them.[7] In the motor group

[6] D. E. McAllister, The effects of various kinds of relevant verbal pretraining on subsequent motor performance, *J. exp. Psychol.*, 46, 1953, 329-336.

[7] J. A. Glaze, The association value of non-sense syllables, *J. genet. Psychol.*, 35, 1928, 255-269.

the first three fingers of each hand were used, and a cover was placed over the keys that *S* could see neither the keys nor his fingers. With this design, referred to elsewhere as Design III,[8] there is no way to determine the absolute amount of transfer, but the verbal and motor groups can be compared to determine which group shows more transfer.

The first task used the same self-paced forced-anticipation method as Experiment II. Three different orders of presentation were used, the inter-trial interval was 2 sec., and learning was to a criterion of two consecutive perfect trials. The transfer-task used a method of paced anticipation with a 2-sec. rate (2 sec. for S, 2 sec. for S—R). The inter-trial interval was 5 sec., and learning was to a criterion of one perfect trial. There were 30 *Ss* from the same population as before, and they were randomly assigned to two groups of 15 each.

Results. With one exception the results for the learning of the first tasks were identical with those of Experiment II. The motor responses were significantly harder to learn, and the difference between motor and verbal learning increased the higher the criterion of mastery. The one exception was the average time per trial; in this experiment the motor group took more time per trial than the verbal group, but this difference was not statistically significant.

TABLE II

MEDIAN TRIALS TO SUCCESSIVE CRITERIA ON TEST OF TRANSFER AFTER PREVIOUSLY LEARNING TASK WITH MOTOR OR VERBAL RESPONSES

Criterion	Motor	Verbal	Difference
1/6	2	2	0
2/6	3	5	−2
3/6	6	5	1
4/6	9	8	1
5/6	18	13	5
6/6	34	24	10

On the transfer-task the motor group was more variable than the verbal group in terms of trials to criterion. To test the significance of this difference the Mann-Whitney *U* test was used, and as predicted, the motor group was found to require significantly more trials to reach criterion than the verbal group ($P < 0.05$, one tail). Table II shows the median number of trials to reach each successive criterion of mastery, and it is apparent that the difference between groups increased with the criterion of mastery.

It should be remembered that the two groups learned exactly the same task (Task 2); the only difference between them was in Task 1. Also, these results differ from those of many studies of transfer in that here the transfer was most pronounced at the end of learning, not at the beginning.

[8] Murdock, Transfer designs and formulas, *Psychol. Bull.*, 54, 1957, 313-326.

EXPERIMENT IV

One explanation of the superiority of verbal over motor responses in both learning and transfer is that cues produced by verbal responses are more distinctive than cues from motor responses. If this is so, then it should be easier to learn a verbal task than a comparable motor task, both involving a serial performance. In a serial task the stimuli for any given response (except the first) are the cues produced by the previous response, and presumably the more distinctive these cues, the easier the learning. The purpose of Experiment IV was to determine if, as predicted, a serial verbal task is easier to learn than a serial motor task.

Procedure. Each *S* learned two tasks, a serial verbal task and a serial motor task. The serial verbal task consisted of 10 nonsense-syllables: WIS, KEN, FAM, REG, DIF, GUL, PED, NOV, VAC, HOB.[9] In the motor task *S* learned the order in which to depress 10 telegraph keys. For both tasks alternate training and testing trials were used. The correct order was indicated to *S* by means of flash cards for the verbal task, and by depressing the keys as *S* rested his fingers on them for the motor task. Then, *S* attempted to duplicate this order himself, either by calling aloud the syllables or by pressing the keys in order. The motor task differs from many others in that the responses were simple, required little skill, and (perhaps most important) had only the cues from the preceding response as stimuli.

The training trials for both tasks used a 2-sec. rate, but all testing trials were self-paced. Learning was to a criterion of two consecutive perfect trials. The order in which *Ss* learned the two tasks was counter-balanced, half learning the verbal first and the other half the motor first. A simple and brief familiarization preceded each task. For the verbal task the 10 cards with the nonsense-syllables were spread out in random fashion on the table and *S* was given 60 sec. to study them. For the motor task *S* pressed each of the keys several times to get the feel of them.

The serial order for each of the two tasks was randomly determined subject to the following restrictions. For the verbal orders, no two adjacent syllables could have the same first, middle, or last letter. For the motor orders, no two adjacent keys could come in succession and regular runs of three or more (*e.g.* 1,3,5) were excluded. There were 32 *Ss*, all from the same population as before.

Results. To reach criterion the verbal task required a mean of 6.9 trials and the motor task a mean of 11.1 trials, and this difference was statistically significant ($t = 4.89$, $P < 0.01$). The median number of trials required to reach each successive criterion is shown in Table III, and as before the difference between verbal and motor tasks increased with higher levels of mastery.

Since on the average the second task for each *S* required two trials less than the first task, a correction factor was used in the following analysis. For each *S* two trials were added to the actual number required for the sec-

[9] Murdock, *op. cit.*, *J. exp. Psychol.*, 55, 1958, 170.

ond task, half the cases representing verbal learning and half motor. By the test for correlated variances the motor task was significantly more variable than the verbal task ($t = 6.23$, $P < 0.01$). (This does not bias the t-test for the difference between verbal and motor tasks reported above, since this test used matched scores.) The women did better than the men on the motor task ($t = 2.18$, $P < 0.05$), but there was no sex difference on the verbal task. The coefficient of correlation between scores on the *ACE* and number of trials to criterion on the verbal and motor tasks were −0.58 and −0.54, respectively (for both, $P < 0.01$). Finally, on the motor task approximately half the *S*s reported using a system (*i.e.* numbering either

TABLE III

Median Trials to Successive Criteria on Serial Learning of Lists Employing Motor or Verbal Responses

Criterion	Motor	Verbal	Difference
1/10	1	1	0
2/10	1	1	0
3/10	1	1	0
4/10	2	2	0
5/10	3	2	1
6/10	4	3	1
7/10	6	4	2
8/10	6	4	2
9/10	7	5	2
10/10	7.5	5	2.5
2×10	10	6	4

fingers or keys), but the means of those that did and those that did not use a system differed slightly and not significantly.

The results of this experiment are consistent with the hypothesis that cues from a verbal response are more distinctive than cues from a motor response, though obviously other explanations for these results could be suggested. Also, the specific verbal and motor responses used in Experiment IV are like the responses used in other experiments of this series, but clearly there is no guarantee that these results could be generalized to still other verbal and motor responses.

Experiment V

The purpose of Experiment V was to determine if there would be a difference in transfer when the first task had a high degree of intra-list similarity as contrasted with low intra-list similarity. Two levels were tested, high (*HS*) and low (*LS*).

Procedure. A recent experiment suggested that strength of association may be a more effective way of measuring similarity than synonymity of meaning.[10] Strength

[10] C. N. Cofer and M. Yarczower, Further study of implicit verbal chaining in paired-associate learning, *Psychol. Rep.*, 3, 1957, 453-456.

of association was used, therefore, to adjust the difference in similarity among the response-terms. There were four response-terms in each list. For the *HS*-condition adjectives were chosen which, according to the Kent-Rosanoff manual, were common associations to the word DARK (and to each other); they were LIGHT, BLACK, and WHITE.[11] In the *LS*-condition the word DARK was again used, but this time the three other terms were adjectives which, according to the same manual, were not common associations to it (or to each other); they were SOUR, COLD, and LONG.

Each *S* served as his own control.[12] After learning either *HS* or *LS* verbal responses to the four stimulus-conditions of Task 1, *S* learned discriminative

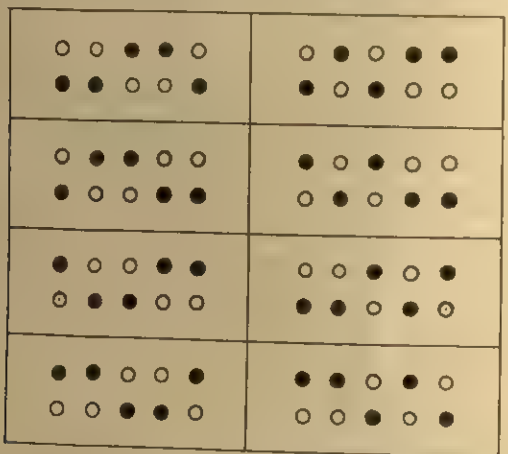


FIG. 4. STIMULUS-CONDITIONS IN EXPERIMENT V
A filled circle represents a light on, an unfilled circle a light off.

motor responses to eight stimulus-conditions of Task 2. Four of these eight had been used on the first task, and these were the experimental (*EX*) conditions; the four new conditions were the control (*C*). The *EX* and *C* conditions were reversed for half the *Ss*.

In all, eight conditions were required, two sets of four each. The principles determining the selection of these sets have been stated elsewhere.[13] The stimulus-conditions are shown in Fig. 4, one set in each column. On Task 2, *S* had to learn the correct channel of the Star-D for each of the eight conditions.

Both tasks used a varied order of presentation and a paced-anticipation method. The presentation rate was 2 sec. (2 sec. for *S*, 2 sec. for S—R) throughout. On Task 1, the inter-trial interval was 6 sec., and the task was learned to a criterion of three consecutive perfect trials. On Task 2, *Ex* and *C* pairs alternated, and after every 4 pairs there was a 6-sec. interval. One set of stimulus-conditions was paired

[11] G. H. Kent and A. J. Rosanoff, A study of association in insanity, *Amer. J. Insanity*, 67, 1910, 37-96, 317-320.
[12] Murdock, *op. cit.*, *Psychol. Bull.*, 54, 1957, 315.
[13] Murdock, *op. cit.*, *J. exp. Psychol.*, 55, 1958, 168.

with the channels in the 90°, 180°, 270°, and 360° positions, and the other set was paired with the channels in the 45°, 135°, 225°, and 315° positions; *S* was informed of this in the instructions. On Task 2, all *Ss* learned to a criterion of two consecutive perfect trials (where a trial consists of one presentation of all eight pairs) or for 24 trials, which ever came first. There were 24 *Ss* from the same population, and they were randomly assigned to the two groups.

Results. The measure of transfer was given by the formula, $T \doteq [(Ex - C)/(Ex + C)]\cdot 100$, where *E* and *C* are now the number of correct anticipations of *Ex* and *C* pairs, respectively. For the *LS* group the mean transfer was + 14.1% ($SD = 31.5\%$), and for the *HS* group the mean transfer was − 12.3% ($SD = 26.2\%$); the difference between the two means was significant ($t = 2.13$, $P < 0.05$). On the first task, the *LS* group unexpectedly required more trials to reach criterion than the *HS* group (means of 45.7 and 31.0 trials, respectively), but this difference was not significant ($t = 1.79, P > 0.05$).

The main results of this experiment were positive; as predicted, the group that learned the *HS* responses on Task 1 showed less transfer than the group that learned the *LS* responses on Task 1.[14]

Discussion

Experiments II and III substantiate a previous finding that motor responses are harder to learn than verbal responses and, once learned, result in less transfer to a second task. Experiment I rules out the possibility that the differences in learning reflect the fact that motor responses require more time to execute than the verbal responses. Experiment V shows that verbal responses with high intra-list similarity result in less transfer to a second task than verbal responses with low intra-list similarity. Experiment IV provides some support for the assumption that cues from a motor response are less distinctive than cues produced by a verbal response.

It seems unlikely that the greater transfer from a verbal response on the one hand, and from a *LS* verbal response on the other hand, could be accounted for in terms of any of the following possible explanations.

(a) Warm-up and learning-how-to-learn. In the first case (*i.e.* verbal and motor responses), the motor group required more trials to reach criterion on Task 1, and thus should demonstrate, if anything, more warm-up and a better-developed learning set. In the case of *HS*- and *LS*-verbal responses, the design was such that warm-up and learning-how-to-learn were equated.

[14] A replication of this experiment using nonsense-syllables instead of adjectives as labels failed to find differences in transfer as a function of intra-list response similarity. This is consistent with the results of J. R. Gerjuoy (Discrimination learning as a function of the similarity of the stimulus names, Unpublished Doctoral dissertation, State University of Iowa, 1953) who also found no differences in transfer with nonsense syllables.

(b) Attention to cues. Attention to the stimulus-patterns and active search for identifying features would presumably occur irrespective of the specific responses required. Actually, one might almost expect more attention to cues with the more similar responses because of the greater difficulty of the task, yet the more similar responses in both cases resulted in less transfer.

(c) Activation of preëxperimentally acquired names. Probably the stimulus-patterns were sufficiently novel that no names had been acquired prior to the start of the experiment; our *S*s could have supplied their own names for the patterns. This would be more likely for the similar responses, so clearly the possibility of names supplied by *S* cannot explain the results.

(d) Meaningfulness. It could be argued that attaching a verbal label to a relatively meaningless stimulus-condition makes that condition more meaningful than if the response were a motor response. This explanation, however, does not seem to apply in the second case. By themselves, the adjectives DARK, LIGHT, BLACK, and WHITE would not seem to make stimuli any more or any less meaningful than the adjectives DARK, SOUR, COLD, and LONG. That is, a *LS* adjective could make its stimulus-condition more meaningful but an *HS* adjective would do the same; there is no reason to assume that the increase in meaningfulness would be greater for one than the other.

(e) Reduction of intra-list generalization. In both cases the stimulus-conditions for the two types of responses were the same, and the two types of responses were learned to the same criterion. Therefore, in both cases generalization-tendencies would be reduced, but the reduction should be the same for both types of responses.

The concept of the acquired distinctiveness of cues would appear to provide one explanation for the results obtained.[15] According to this concept, attaching dissimilar responses to similar stimulus-cues should increase the distinctiveness of the cues and so facilitate new learning. On this basis it would be expected that verbal responses and *LS*-responses would mediate new learning more effectively than motor responses and *HS*-responses, respectively. This is the interpretation suggested by Goss and Greenfeld, who found that, with sufficient training (overlearning) on Task 1, verbal responses resulted in more transfer than other types of responses.[16]

It is true that several studies have failed to find differences in transfer as a function of differences in the distinctiveness of the cues.[17] One possible explanation of why some investigators' studies find differences and others

[15] John Dollard and N. E. Miller, *Personality and Psychotherapy*, 1950, 98-105.

[16] A. E. Goss and Norman Greenfeld, Transfer to a motor task as influenced by conditions and degree of prior discrimination training, *J. exp. Psychol.*, 55, 1958, 258-269.

[17] M. D. Arnoult, Transfer of predifferentiation training in simple and multiple shape discrimination, *J. exp. Psychol.*, 45, 1953, 401-409; H. W. Hake and C. E. Eriksen, Effect of number of permissible response categories on learning of a constant number of visual stimuli, *ibid.*, 50, 1955, 161-167; J. S. Robinson, The effect of learning verbal labels for stimuli on their later discrimination, *ibid.*, 49, 1955, 112-114.

do not is rather evident from a comparison of the studies. For the most part, those studies reporting differences continue practice on the first task well beyond one perfect trial, while those studies that fail to find differences do not carry practice much beyond one perfect trial, if that far. It may be that a high level of proficiency is necessary for responses learned on the first task to function as mediating responses on the second task, and that greater transfer as a function of greater distinctiveness of cues will only be found when a sufficiently high level of distinctiveness is attained.

Experiment VI

The previous experiments are consistent with the hypothesis that the responses learned in Task 1 mediate the learning of Task 2. They suggest that mediating responses are important in transfer. However, as Osgood has pointed out,[18] despite the wide use of mediating responses as an explanatory concept, there have been very few studies which have observed and measured the mediating response directly. The emphasis is usually on the *mediated* response rather than on the mediating response. Experiment VI is an attempt to obtain direct evidence on certain characteristics of mediating responses.

One situation in which mediating responses are presumed to play an important role is in the development of instrumental sequences, or chaining.[19] For instance, tying a necktie, reciting the alphabet, or learning to play a particular passage on the piano would be familiar examples. In the development of such a sequence it is hypothesized that initially the individual responses are initiated one at a time by external conditions; with practice the external conditions become less important, the cues from the preceding response function as stimuli, and the skill runs itself as it were. Each response then is a cue-producing or mediating response whose response-produced stimulation is one link in the chain.

While this seems eminently reasonable, there seems to be very little direct evidence that chaining actually develops in such a fashion. While the development of skills has been widely studied, usually there is not the detailed analysis of the individual components of the skill that this theory requires. The task in Experiment VI consisted, accordingly, of a simple chaining situation in which *S* had to perfect a motor skill that involved pressing eight telegraph keys in a specified sequence. This task was selected because the supposed mediating responses are overt and easily recorded, hence can be observed rather than inferred.

---

[18] C. E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 411.
[19] Osgood, *ibid.*, 399-401.

Procedure. The telegraph keys were designated 1 to 8 from left to right, and a card with the sequence (*e.g.* 4 8 2 5 1 7 3 6) was always before *S*. Thus, *S* did not have to learn the serial order. Each *S* used the four fingers of each hand but not the thumb; the keys were enclosed in a box with a curtain so *S* could see neither the keys nor his hands. All *S*s were given 50 trials on one sequence, the first 20 trials spaced 30 sec. apart, the next 20 trials spaced 10 sec. apart, and the last 10 trials spaced 30 sec. apart.

With this task, initially each reaction should be made to the particular number on the card, but with practice the sequence should become automatic. Each response in the sequence (except the last) should come to function as a mediating response and cue off the following response. To study the individual responses the time at which each response occurred was recorded by 1/100-sec. Standard Electric Timers. Thus, the latency of each response was measured, where the latency was the time between a given response and the preceding response.

There were 16 *S*s from the same population as before, and each *S* practiced on one sequence. In all there were eight different sequences, and a Latin square was used to counter-balance the serial position of each key. Also, for each *S* the simple reaction-time to a proprioceptive stimulus was measured; specifically, this was the time required to lift the right hand from a telegraph key after feeling an upward movement in the left hand.

Results. On each trial the latency of each response (except the first) was determined, and the median of these seven latencies was the value for that trial. An examination of the results indicated that the latencies decreased markedly with practice, that these latencies appeared to be approaching a lower limit, and that, for many *S*s, this lower limit was very close to the proprioceptive *RT*.

The following procedure was adopted to test this statistically: the latency curve was assumed to be an exponential function of log trials which approached the proprioceptive *RT* as a limit. For each *S* individually the best-fitting curve was determined on the basis of only the first 10 trials (in blocks of two) and the proprioceptive *RT*. For each *S* individually this made it possible to predict the value for latency on Trial 45, and this was then compared with the obtained latency on Trial 45 (which was considered the median of Trials 41–50). For all 16 *S*s the mean difference between the predicted and obtained values was 20 m.sec., which is rather small considering that errors of measurement in the apparatus itself are at least of the order of 10 m.sec. Also, the difference between predicted and obtained values did not begin to approach statistical significance ($t = 0.34$).

This analysis clearly suggests that the latencies of the individual responses in this chaining situation were approaching the proprioceptive *RT* as a lower limit, as a prediction based on this assumption was quite accurate. Although this specific result has not been predicted in so many words by a mediation theory, still it is obviously consistent with the notion that stimu-

lation from each response serves to cue off the subsequent response. Also, one objection to a mediation hypothesis points out that the individual components of a highly skilled act (such as a pianist playing a rapid cadenza) occur too rapidly to be cued off by preceding responses, hence must be centrally determined.[20] While the results presented here certainly do not prove that responses cannot occur this rapidly, still there is nothing in these results to prove that they do occur this rapidly.

SUMMARY

Six experiments were conducted to investigate further the role of response-factors in learning and transfer. Experiment I showed that the difference between learning verbal and motor responses could not be attributed to a longer time required to execute the motor responses, as the motor responses had a consistently shorter reaction-time than verbal responses. Experiments II and III verified the previous findings that learning and transfer are facilitated with verbal responses. The conditions differed somewhat from the earlier experiments. Experiment IV provided some support for the hypothesis that cues produced by verbal responses are more distinctive than cues from motor responses by showing that a verbal serial list was easier to learn than a motor serial list. Experiment V showed that labels (adjectives) of low intra-list similarity resulted in more transfer than labels of high intra-list similarity. These results were interpreted as supporting the concept of acquired distinctiveness of cues, and it appeared that responses learned in Task 1 served to mediate Task 2 learning. Experiment VI was conducted to obtain direct evidence on mediating responses, and the results of a study of chaining supported a mediation hypothesis by showing that the latencies of the individual responses appeared to approach the proprioceptive reaction time as a limit.

[20] Osgood, *ibid.*, 401.

FURTHER EVIDENCE FOR THE INTERACTION OF DRIVE AND REWARD

By JOHN P. SEWARD, RICHARD A. SHEA, and ROBERT H. DAVENPORT
University of California, Los Angeles

In an earlier study we measured the effect of two deprivation-intervals (0 and 22½ hr.) and two amounts of reward (0 and ½ or 1 gm.) on speed in a runway and found a highly significant interaction between them.[1] This result was interpreted as evidence in favor of Hull's assumption that drive (*D*) and incentive (*K*) multiply in determining reaction-potential,[2] and against Spence's view that they add.[3] Varying delay of reinforcement (1 and 5 sec.) rather than amount, Loess found no significant difference between deprivation-intervals of 3 and 22 hr. in their effect on speed of response.[4] His finding was confirmed by Ramond, using 4-hr. and 22-hr. intervals, and by Besch and Reynolds, varying length of alley.[5]

Our results might be predicted, as Loess has pointed out, on the ground that, since a reward of zero means no reinforcement, we really varied habit-strength; this would account for its interaction with drive.[6] The same argument would not, however, apply to Behan's experiment.[7] Behan trained rats to push a panel for large or small rewards of food or water, and tested them after six deprivation-intervals from 6 to 48 hr. Both level of activity and force of response showed significant interaction between deprivation-interval and amount of reward. It is possible, therefore, that the magnitude of a reinforcer interacts with drive, but that delay does not. The disagreement may also be due to the different ranges covered or the measures used.

---

* Received for publication July 20, 1959. This research was aided by a grant from the National Science Foundation.

[1] J. P. Seward, R. A. Shea, and David Elkind, Evidence for the interaction of drive and reward, this JOURNAL, 71, 1958, 404-407.

[2] C. L. Hull, *A Behavior System*, 1952, 6-7.

[3] K. W. Spence, *Behavior Theory and Conditioning*, 1956, 197 f.

[4] H. B. Loess, The effect of motivational level and changes in motivational level on performance in learning, Unpublished Doctoral dissertation, State University of Iowa, 1952.

[5] C. K. Ramond, Performance in instrumental learning as a joint function of delay of reinforcement and time of deprivation, *J. exp. Psychol.*, 47, 1954, 248-250; N. F. Besch and W. F. Reynolds, Alley length and time of food deprivation in instrumental reward learning, *ibid.*, 56, 1958, 448-452.

[6] Loess, *op. cit.*, 33. Spence himself explicitly rejected this position (*op. cit.*, 218).

[7] R. A. Behan, The quantification of drive: III. Privation of food and water, Unpublished Doctoral dissertation, Michigan State College, 1954.

Significant interactions may appear only over a wide range or only within certain intervals.

Experiment I

In the present study, we increased the number of values of drive and reward and attempted to measure their joint effect on running speed. The drive was hunger, and the reward was food. In a second study, also reported here, we used thirst and water.

Apparatus. The enclosed runway, made of wood and covered with hardware-cloth, was 11 ft. long, 3 in. wide, and 4 in. high. A start-box, 7½ × 3 × 4 in., was equipped with two sliding doors, the first opaque, the second transparent. A bank of end-boxes, each 16 × 4 × 4 in., mounted on rollers made it possible to run one *S* while others were eating. Seven inches from the start and end of the runway were two photocells connected with 0.01-sec. and 0.1-sec. electric timers, respectively. All wooden parts were painted gray. Curtains of cheesecloth on both sides and shaded overhead lights behind diffusing glass approximated one-way screening.

TABLE I

Treatments and Number of Ss Assigned to Each in Experiment I

Amount of reward	Percentage of intake prefed						N
	100		50		1		
Balance	*a**	11	*b*	10	*c*	7	28
1%	*d*	11	*e*	10	*f*	7	28
0	*g*	11	*h*	10	*i*	7	28
N		33		30		21	84

* Letters designate treatments for later reference.

Subjects. We used 84 male albino rats, 3–4 mo. old at the outset of the work.

Design. Drive-strength was defined by the amount of the daily ration prefed; reward-value, by the amount fed in the goal-box. There were three levels of drive—1%, 50%, and 100% of daily intake prefed—and three levels of reward—0, 1%, and the balance of the day's ration.[8]

The *Ss* were studied in five replications of 10–23 each. Within the first three replications they were assigned at random to one of three amounts of prefeeding and to one of two amounts of reward—1% or the balance. In the fourth replication, all *Ss* ran to no reward. The last replication served to make the subgroups proportional between the main classes.[9] The nine treatments used, with the number of *Ss* assigned to each, are shown in Table I.

Maintenance. All *Ss* were housed four in a cage with water but no food. Once a

[8] The percentages were defined as follows: 1% meant two pellets totaling 0.1–0.2 gm., actually about 0.5–1% of *S*'s daily intake; 50% meant half of *S*'s median intake for the last 16 days before training; 100% meant as much as *S* would eat from a total of 24 gm; 'balance' meant the remainder from a 24-gm. total (except in the first and last replications, when the groups prefed 1% and 50% received the remainder from their median intake).

[9] This failure to randomize replications completely was unfortunate but unavoidable, since the condition of nonreward was added late.

day they were fed small pellets for 1 hr. in individual feeding cages with water available. Daily consumption was measured by weighing each *S*'s ration before and after. This regime continued during the adaptation described below and thereafter, until *S* reached a 'stable level' of food-intake. Stability was ascertained by finding the median intake for 16 days. The criterion was met when the number of scores above and below this median did not differ by more than two for both the first and last eight days. The *Ss* usually were maintained on this schedule about five weeks before training started.

Adaptation. Two or three weeks after starting their feeding cycle, *Ss* were adapted to the apparatus. This was done in several stages: Two days of group exploration, 10 min. a day, were followed by two days of individual exploration, 5 min. a day. At this time *S* was assigned to a small black detention-box and put in it on both days for 5 min. at a time. Then came one timed trial a day to an end-box containing water but no food, from which *S* was transferred to the detention-box for periods of confinement increasing progressively from 10–25 min. These trials took place after feeding in Replications 1 to 3, before feeding in Replications 4 and 5. No drinking was observed in the end-box. The trials varied in number from a minimum of 8 to about 12, depending on how long it took *S* to reach the criterion of food-intake. During all stages of adaptation, the end-boxes were shifted frequently to provide general familiarity.

Training. The day after an adapted *S* met the feeding criterion his training began. It consisted of one trial a day conducted as follows: *S* first was prefed in an individual feeding cage and the end-box supplied with water and baited or not, as required. When *S* had finished prefeeding, it was put in the start-box. As it faced the doors, *E* raised the opaque door, counted to three, and raised the transparent door, automatically starting both timers. If *S* failed to stop the second timer within 3 min. it was urged; after five successive failures, its training was discontinued. After eating the reward in the end-box, if it had already been given its full ration, *S* was returned directly to the living cage. If not, it was put first in its detention-box for 30 min., then in its feeding cage for the rest of its ration,[10] defined as its median intake on reaching stability.

Two times were recorded: starting time, from raising the second door till *S* reached the first photocell; and running time, from the first photocell to the second. Training continued for each *S* until its performance 'leveled off,' *i.e.* until both measures had met a criterion of stability computed as follows: The median was determined for 16 consecutive days; both for the first 8 days and for the last 8 days the difference between the number of scores above and below that median had to be no greater than 1.

Results. Each *S* contributed one starting time and one running time to the data, the median of the 16 days on which it met the criterion of stability for that measure. Individual scores were first converted into reciprocals and multiplied by 100. Reciprocals were more suitable than time scores for analysis of variance: (1) because their distributions were less skewed and

[10] This was an error that was detected too late to rectify. All groups should have spent 30 min. in detention-boxes before returning to their living cages. A control-experiment has since been done and the error shown to be without effect.

more homogeneous in variance; and (2) because *S*s discarded for 5 consecutive runs of over 180 sec. could be scored 1 and included with little distortion (since the reciprocal of infinite time would be zero).

Mean running speeds for the various motivating conditions are given in Table II. The three classes of reward show linear, slightly divergent, upward trends with decrease in prefeeding. It is noteworthy that the means for 1% reward fall nearly midway between those for zero and 'balance.'

Although Bartlett's test for heterogeneous variance was significant at the 5% level, we decided to use the *F*-test with the usual reservations.[11] Both main effects proved to be highly significant ($p < 0.001$) and their interaction reached the 5% level. We therefore applied *t*-tests to the differences between individual prefeeding groups and between differently rewarded groups. The results appear in Tables III and IV, respectively. Table III shows that prefeeding groups differ most significantly when run to the larger reward. It is of interest to note, however, that the difference between groups prefed 1% and 100% is significant even with no reward in the end-box. Table IV shows that amount of reward produced the most significant differences following the least prefeeding.[12]

TABLE II
MEAN RUNNING SPEEDS (RUNNING TIME RECIPROCALS×100) IN EXPERIMENT I

Amount of reward	Percentage of intake prefed			Mean
	100	50	1	
Balance	7.6	20.2	31.6	18.1
1%	6.9	13.6	21.0	11.8
0	4.4	8.4	12.7	7.9
Mean	6.3	14.0	21.8	12.9

It was important for our problem to locate the source of the interaction, which we did by breaking down our 3 × 3 table into its nine possible 2 × 2 tables and finding *t* for the interaction within each. Three *t*s were significant. These can be identified, by reference to Table I, as belonging to the treatments labeled *acgi* ($p = 0.005$), *abgh* ($p = 0.05$), and *acdf* ($p = 0.05$). Note that the first two involve a reward of zero, but the third does not.

In general, starting speeds showed the same trends as running speeds. The main effects of both prefeeding and reward were significant at the

[11] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1953, 85f.

[12] In spite of heterogeneity, the mean-square for error from the analysis of variance was used in computing all *t*s. This was done to make more comparable the *t*s for different parts of Table II.

1/10% level. Their interaction was not significant, however, and the variance within groups was high, and therefore we did not analyze the data further.

Experiment II

In our second experiment we hoped to elaborate the first by tapping more levels of drive. We changed to thirst for several reasons. One was the effort or expense involved in getting uniform pellets for accurate weighing of small quantities. Another was the length of time required for rats to stabilize their food-intake.[13] On the positive side, we were interested in the generality of drive-reward interaction.[14]

TABLE III

VALUES OF *t* FOR DIFFERENCES BETWEEN GROUPS PREFED DIFFERENT AMOUNTS IN EXPERIMENT I

Amount of reward	Percentage of intake prefed		
	1–50	1–100	50–100
Balance	3.33*	7.16*	4.16*
1%	2.16†	4.21*	2.21†
0	1.26	2.48†	1.32

* Significant at the 1% level.
† Significant at the 5% level.

TABLE IV

VALUES OF *t* FOR DIFFERENCES BETWEEN GROUPS RUN TO DIFFERENT AMOUNTS OF REWARD IN EXPERIMENT I

Amount of reward	Percentage of intake prefed		
	100	50	1
Balance—1%	0.24	2.13†	2.86*
Balance—0	1.08	3.81*	5.11*
1%—0	0.85	1.68	2.24†

* Significant at the 1% level.
† Significant at the 5% level.

Apparatus. The same apparatus was used as in Experiment I.

Subjects. Data were based on 120 male albino rats, 3–4 mo. old at the start of the experiment.

Design. Again drive and reward were manipulated by varying the amounts of incentive-object (in this case water) provided just before and after the measured response. Fifteen treatments combined five levels of predrinking and three of re-

[13] A third problem, intriguing in itself, arose in connection with the group run to 'balance' after 100% prefeeding. Some *Ss* developed a tendency to eat less and less of their ration in the feeding cage and more of it in the end-box.

[14] Behan (*op. cit.*) reported a significant second-order interaction involving hours of deprivation, amount of reward, and type of motive (hungry-to-food *versus* thirsty-to-water).

ward, as shown in Table V.[15] The Ss were run in 8 replications of 12–16 Ss each. Partial randomization was achieved. Ideally, each treatment should have been represented once in each replication. This was not done, partly due to scheduling difficulties and loss of Ss during adaptation, but mainly because the condition of zero-reward was added after the experiment was almost half over. The first three replications and the fifth and seventh covered eight treatments each and contained two Ss per treatment with few exceptions. Replications 4, 6, and 8 were run to zero-reward 2–4 Ss assigned to each of the five amounts of predrinking. As in Experiment I, effects of treatment were partly confounded with a possible effect of replications, though no such effect actually was observed.

Maintenance. The Ss were housed four in a cage with large pellets but no water. Once a day they were given water for 45 min. (for 1 hr. on the first three days)

TABLE V

NUMBER OF Ss ASSIGNED TO EACH TREATMENT IN EXPERIMENT II

Amount of reward	Percentage of intake predrunk					
	100	67	50	33	1	*N*
Balance	8	8	8	8	8	40
1%	8	8	8	8	8	40
0	8	8	8	8	8	40
Total *Ss*	24	24	24	24	24	120

in individual cages also supplied with large pellets. Intake of water was measured by weighing the water before and after drinking, with an adjustment for dripping and evaporation measured by a control-bottle in an empty cage. After adaptation and before training each *S* had to reach a stable level of water-consumption, defined by the same 16-day criterion used for food. This period lasted 5–6 weeks on the average, with a range from 3–8.

Adaptation. During adaptation, which duplicated the adaptation of Experiment I, the end-boxes contained large pellets only. All Ss were adapted before their daily watering. No eating was observed.

Training. Procedure in training paralleled that for hunger. The end-box was supplied with large pellets and, except in the case of the nonrewarded groups, with the required amount of water in an inverted bottle with nozzle projecting downward through the wire top as in the drinking cages. Each trial consisted, as before, of the following steps: predrinking; the run; drinking of the reward, if any, in the end-box; return to the living cage for groups that had finished their ration, and 30 min. in the detention-box for the others (see footnote 11). The latter had two more steps: more water in the drinking cage; return to the living cage. Except for the condition of 100%-to-balance already noted, all groups received as a total ration the median intake for the last 16 days before training.

Results. Scores were treated as in Experiment I. Each *S*'s 'asymptotic' running time was converted to its reciprocal and multiplied by 100. Means

[15] Percentages of intake were based on the median daily consumption of water for the last 16 days before training, with two exceptions: (a) 1% was approximated by giving 0.2–0.3 cc.; (b) the 100%-to-balance group was given 35 cc before and the remainder in the end-box; amounts actually drunk were recorded.

were then computed for all groups, Bartlett's test satisfied, and an analysis of variance performed. Predrinking was significant beyond the 1/10% level; reward reached the 5% level, as did the interaction.

When we applied t-tests to pairs of treatments, however, we found significant differences due to predrinking only if one amount was less than 50% and the other was 50% or greater. Differences due to reward were significant only when predrinking was less than 50%. Since there were no significant differences within the three largest percentages of predrinking,

TABLE VI

MEAN RUNNING SPEEDS (GROUPS PREDRINKING 50–100% COMBINED) IN EXPERIMENT II

Amount of reward	Percentage of intake predrunk			
	50+	33	1	Mean
Balance	11.0	30.4	38.1	20.3
1%	12.5	24.0	28.9	18.1
0	12.1	17.6	19.8	14.8
Mean	11.9	24.0	28.9	17.7

TABLE VII

VALUES OF t FOR DIFFERENCES BETWEEN GROUPS PREDRINKING DIFFERENT AMOUNTS IN EXPERIMENT II

Amount of reward	Percentages of intake predrunk		
	1–33	1–50	33–50
Balance	1.64	7.08	5.07*
1%	1.04	4.28	3.00*
0	0.47	2.01†	1.44

* Significant at the 1% level.
† Significant at the 5% level.

it was possible to simplify the picture by pooling the three groups run to each amount of reward. Reference to Table V shows that the outcome would be a 3 × 3 table with 24 *S*s in each cell of the first column and 8 in each cell of the other two. The results of this procedure appear in Table VI. Again Bartlett's test was satisfied, and a new analysis of variance gave *F*s for predrinking and reward significant at the 1/10% and 5% levels, as before. Interaction was now significant at the 1% level.

Tables VII and VIII show the results of t-tests applied to differences in predrinking and differences in reward. On the whole, the trends in these tables are comparable with those for food in Tables III and IV. The levels of significance in Table VIII, however, are much lower than those in Table IV, though it should be noted that the ts for 1% predrunk barely miss the 1.98 needed for a p of 0.05.

With Experiment I as a model, we now applied t-tests to the simple interactions in Table VI. The results were closely analogous with those for food; only three interactions claimed attention and they were the same ones that were significant in the first experiment. Treatments *acgi* (identified by Table I) interacted significantly at the 1/10% level and *abgh* fell just short of 1%. The third cluster, *acdf*, critical in that it did not involve absence of reward, gave $t = 1.97$, with 1.98 needed for the 5% level.

In general, starting speed showed the same trends as running speed, but

TABLE VIII

VALUES OF t FOR DIFFERENCES BETWEEN GROUPS RUN TO DIFFERENT AMOUNTS OF REWARD IN EXPERIMENT II

Amount of reward	Percentage of intake predrunk		
	50+	33	1
Balance—1%	−0.55	1.36	1.96
Balance—0	−0.41	2.73*	3.90*
1%—0	0.15	1.36	1.94

* Significant at the 1% level.

with more irregularity and error-variance. Both main effects were highly significant ($p = 0.001$), but their interaction ($p > 0.20$) did not justify further treatment.

DISCUSSION

Problem of interaction. The data of Experiment I suggest that the effect of prefeeding on running speed is to some extent a function of amount of reward. This is true even when the condition of nonreward is disregarded, although two tiny pellets exert an influence out of all proportion to their size. Assuming that these two variables are related to drive and incentive, respectively, we have grounds for asserting that D and K are not altogether independent in their effects on performance. Unfortunately, the statistical basis of the assertion does not command complete confidence.

Experiment II also suggests the presence of interaction between drive and reward, although of border-line significance unless the reward is reduced to zero. Taken together the two experiments increase our confidence in the finding and give it some degree of generality.

The data confirm our suspicion that the existence, or at least the detection, of interaction depends on what values of the independent variables are chosen. In both experiments it was necessary to use a wide range of rewards; in the experiment on thirst only a portion of the predrinking range was critical.

Adding these results to the other studies discussed above, we may con-

clude that the effects of drive and incentive on performance do not merely summate, nor do they merely multiply. More complex relations are likely to emerge as they are measured under a variety of conditions.

Differences between hunger and thirst. Although in general the same, performance motivated by food and by water showed several differences: (1) Running speed to water was apparently less sensitive to variation of reward. (2) Scores under thirst were much more variable than under hunger (mean-squares for error were 87 and 48, respectively). (3) Mean speeds to food and water showed different functional relations to decrease in amount prefed or predrunk. Those to food increased in a linear manner from 100% to 1% prefed, diverging with amount of reward. Those to water showed little trend and no regular divergence between 100% and 50% predrunk. Below that point, the curves to 1% and balance of water-ration started upward, assuming an S-shape. The three differences are probably not coördinate, since the first may well be due to the second and third. An explanation for both of these may be sought in the eating and drinking patterns of rats.

Both Skinner and Bousfield have published fairly smooth, negatively accelerated curves of food consumption plotted against eating-time.[16] Skinner's curves for drinking, however, are much less regular, and he noted that discontinuous curves occurred in thirst but not in hunger.[17] Stellar and Hill found discontinuity characteristic of drinking regardless of deprivation.[18] Their rats drank at maximal rate or not at all, with an 'initial burst' that varied in length with deprivation, but not beyond a limit of about 8 min. for deprivations of 48 hr. or more. Our rats predrank their complete rations in about 10 min. It seems reasonable that after 22-hr. deprivation they would drink about half of the ration in the initial burst. Thereafter they would fluctuate between thirst-related and competing activities. Groups predrinking 50–100% therefore would be expected to perform at a lower and more variable level.

Summary

In Experiment I, 9 groups of rats on a 24-hr. feeding schedule were prefed, 1%, 50%, or 100% of their daily intake and were then trained in a runway to 0, 1%, or the balance of their ration. Results for 'asymptotic'

[16] B. F. Skinner, *The Behavior of Organisms*, 1938, 344-349; W. A. Bousfield, Quantitative indices of the effects of fasting on eating behavior, *J. genet. Psychol.*, 46, 1935, 476-479.

[17] Skinner, *op. cit.*, 358-361.

[18] Eliot Stellar and J. H. Hill, The rat's rate of drinking as a function of water deprivation, *J. comp. physiol. Psychol.*, 45, 1952, 96-102.

running speeds were as follows: (1) Scores increased significantly with decrease in amount prefed and amount of reward, with a significant interaction. (2) The significance of differences between prefeeding groups increased with amount of reward. The significance of differences between reward-groups decreased with amount prefed. (3) Three simple interactions were significant, one of which did not include the condition of nonreward.

In Experiment II, 15 groups of rats on a 24-hr. watering schedule predrank 1%, 33%, 50%, 67%, or 100% of their daily intake and were run to 0, 1%, or the balance of their ration. Results for 'asymptotic' running speeds confirmed those of Experiment I with the following qualifications: (1) For portions of predrinking above 33%, none of the differences between paired groups was significant. (2) The simple interaction, corresponding to that of Experiment I, that excluded nonreward was of borderline significance.

The conclusion drawn from these results is that drive and incentive do not combine by either simple addition or multiplication but in a more complex manner.

BASIS FOR LIGHTNESS-JUDGMENTS OF GRAYS

By RICHARD M. WARREN, National Institute of Mental Health, and E. C. POULTON, Medical Research Council Applied Psychology Research Unit

The *physical-correlate theory* of sensory intensity has been suggested as a general rule governing judgments for all modalities.[1] The theory holds the estimates of sensory intensity are based upon experience with the manner in which sensory excitation is correlated with some physical attribute of the stimulus (not necessarily the one employed conventionally to describe stimulus-intensity). It is considered that the physical correlate upon which visual intensity-judgments are based is distance, *i.e.* an individual's experience with the effect of distance from the light-source on the luminance of a reflecting object. Predictions based upon this theory have been verified recently for brightness-judgments by Warren and Warren,[2] and the present experiment was designed to test these predictions as applied to 'lightness' judgments.

Warren and Warren found that estimates of the decrease in luminance of an object produced by doubling its distance from the light by one group of *Ss* were quantitatively equivalent to judgments of half-brightness by a second group. The authors stressed the point that for estimates of the effect of changing distance from the illuminated field to its light-source to be accurate, contrast-effects with surrounds must be minimized by the use of large contiguous stimulus-fields. With gray papers as well, background-contrast has a considerable effect on 'lightness' judgments under conditions usually employed, hence these judgments depend not only upon the reflectance of the gray stimulus-papers but also upon the reflectance of the surrounds.

In the presence experiment, an attempt was made to minimize the influence of contrast with stimulus-surrounds by having *S* make 'lightness'

* Received for prior publication June 10, 1960. This research was done at the Medical Research Council Applied Psychology Research Unit, Cambridge, England. The research was supported in part by the Air Force Office of Scientific Research of the Air Research and Development Command, United States Air Force, through its European Office.

[1] R. M. Warren, A basis for judgments of sensory intensity, this JOURNAL, 71, 1958, 675-687.

[2] R. M. Warren and R. P. Warren, Basis for judgments of relative brightness, *J. opt. Soc. Amer.*, 48, 1958, 445-450.

judgments of a gray relative to white while viewing large adjacent sheets of gray and white papers. By using only the first 'lightness' judgments, a measure was obtained free from the effects of previous judgments. The advantages of avoiding "prior stimulus context" have been pointed out by Garner.[8]

The quantitative relation predicted from the physical-correlate theory is a simple one. Since we may consider the intensity of light falling on an object as decreasing as the square of its distance from a light-source, it would be predicted that 'lightness' judgments would be proportional to the square root of the reflectance of the gray paper. Since white was called 100, the 'lightness' of any paper should be 10 times the square root of the percent-reflectance relative to white.

Method

Subjects. The 475 *Ss* were men, students of Cambridge University. They were unfamiliar with psychophysical testing.

Material. Sheets of white and of neutral gray papers of different reflectance were purchased from the Mitchell Colour Card Co., London. By cementing sheets of two different grays together back to back and then cutting to the desired size, pairs of comparison-stimuli were produced which could be changed simply by turning over. In some experiments black rayon velvet also was used. The reflectances of all these materials were measured using both a MacBeth Illuminometer and a Salford Electrical Instruments Ltd. Illuminometer.

Procedure. Ss were tested one at a time in the Junior Combination Rooms of several of the Cambridge Colleges. After reading typewritten instructions, *S* was taken to another part of the room and seated before contiguous gray and white stimulus-sheets lying on a wooden table. After *S* wrote down his judgment concerning the 'lightness' of the gray paper relative to white (thinking that only this one judgment was required), the gray sheet was turned over by *E*. With a different gray surface now uppermost, *S* was instructed to make a second 'lightness' judgment. Then *S* was requested not to discuss the experiment with anyone who had not already served as *S*, the gray sheet was turned back to the first side, and the next *S* brought over. A group of 50 different *Ss* were used for each set of stimulus-conditions. In addition to the 400 *Ss* who followed instructions properly, 75 *Ss* did not follow instructions and their data were rejected. The stimulus-display was rotated 180° for half the *Ss* in each group to avoid any right-left position-effects. The ambient illumination was always at a level adequate for comfortable reading.

Experiment I (Numerical Estimates). S saw white and gray rectangular strips, each 11¼ × 8⅞ in., with one of the larger edges of each strip contiguous with the other. (Actually the white sheet was very slightly less than 11¼ in. and extended for some distance beneath the gray.)

Three groups of 50 *Ss* each were employed; the reflectances of the first and second grays presented to each group are given in Table 1. Each *S* read the following in-

[8] W. R. Garner, Half loudness judgments without prior stimulus context, *J. exp. Psychol.*, 55, 1958, 482-485.

structions: "You will see two pieces of paper lying side by side, one white and one gray. If the lightness of the white paper is called 100, what (smaller) number describes the lightness of the gray paper? Please write down your answer on the slip of paper provided."

Experiment II (Linear Estimates). The stimulus-situation was the same as in Experiment I, but *S* was required to indicate the 'lightness' of the gray by placing a mark on a line marked 'white' at one end and 'black' at the other.

One group of 50 *Ss* was employed. As shown in Table 1, they received stimuli of the same reflectance as one of the groups in Experiment I. *Ss* read the following instructions: "You will see two pieces of paper lying side by side, one white and the other gray. You are to judge the lightness or darkness of this gray. A slip will be given you marked 'white' at one end and 'black' at the other. Consider this line to represent a scale of grays lying between white and black. Please place a mark on this line to indicate where you think the gray paper belongs on this scale." The end of the line marked 'white' was placed on either the left or right to agree with the lateral position of the white paper relative to the gray.

Experiment III (Black Surrounds). Pairs of gray paper used for Experiment I were employed, but in this experiment, instead of viewing the gray and the white papers against the wooden table as background, the *Ss* saw them against a black rayon velvet background. The velvet (reflectance = 0.31%) extended 4 in. beyond both top and bottom and 6 in. beyond each of the sides of the large rectangle formed by the gray and white strips.

Two groups of 50 *Ss* each were employed. *Ss* read the following instructions: "You will see two strips of paper, one white and one gray, lying side by side upon some velvet. If the lightness of the white strip is called 100, what (smaller) number describes the lightness of the gray strip? Please write down your answer on the slip of paper provided."

Experiment IV (Asymmetrical Contrast). Three parallel strips were seen by *S*, white, gray, and black rayon velvet. Each strip measured 11¼ × 5⅞ in. The gray was always in the middle, contiguous with both the white and black strips. Two groups of 50 *Ss* were employed, and the same instructions were used as in Experiment III.

Results

The results are summarized in Table 1 for each of the four experiments. The reflectance (relative to white paper = 100) of the first gray paper presented to each group of 50 *Ss* given in the left hand column, then the theoretical value calculated on the basis that 'lightness' of white = 100, followed by the observed median, mean first quartile, and third quartile.[4] The second half of the same line gives the same data concerning the second gray paper presented to the group.

The first judgments of *Ss* were uninfluenced by previous estimates and hence afford a better test of predictions based on the physical-correlate

[4] Due to the response-limits of 0 and 100, the distribution of judgments is skewed for those papers either lighter or darker than middle gray. This asymmetric distribu tion makes the median rather than the mean the statistic of choice.

theory than do the second judgments. Experiment I shows that, with large contiguous stimulus-fields, calculations based upon the hypothesis that 'lightness' is proportional to $10\sqrt{R}$ are in agreement with experimental results.

A median judgment of 50 (half the 'lightness' of white) was obtained for the gray paper having a reflectance of 22.4% relative to white. This

TABLE 1

'LIGHTNESS' JUDGMENTS OF GRAY PAPERS RELATIVE TO A STANDARD WHITE PAPER CALLED 100

First judgment						Second judgment					
Percent reflectance (R)*	$10\sqrt{R}$ (Theoretical Response)	Median	Mean	Q_1	Q_3	Percent reflectance (R)*	$10\sqrt{R}$ (Theoretical Response)	Median	Mean	Q_1	Q_3
Experiment I (Numerical estimates)											
9.43	31	30	29	20	40	53.0	73	75	70	60	80
22.4	47	50	47	37	60	53.0	73	80	76	72	85
53.0	73	65	61	55	74	22.4	47	40	41	30	52
Experiment II (Linear estimates)†											
9.43	31	34	38	27	42	53.0	73	79	78	71	86
Experiment III (Black surrounds)											
22.4	47	50	50	40	65	53.0	73	75	73	65	85
53.0	73	67.5	62	50	75	22.4	47	40	40	30	50
Experiment IV (Asymmetrical contrast)											
9.43	31	45	48	35	60	53.0	73	80	79	73	85
53.0	73	70	67	60	78	22.4	47	49	45	40	50

* The reflectance of the gray papers are expressed as percentage of the reflectance of the standard white paper. The absolute reflectance of this white paper is 85%.

† The numerical values represent the percentages of the total distance from black to white represented by *S*'s mark on the line.

agrees closely with the brightness-fractionation experiment by Warren and Warren in which half subjective-intensity judgments with large contiguous stimulus-fields also approximated 25% of the stimulus intensity.[5]

In Experiment II, results similar to numerical estimates were obtained by having *S* place a mark on a line labelled 'white' at one end and 'black' at the other. It is interesting that this agreement was found in spite of the fact that the mark on the line was an 'interval' judgment on the continuum between white and black, while the numerical estimate was a type of 'ratio' judgment (no mention of zero or black was made).

Experiment III demonstrates that the results obtained for Experiments I and II were not a function of the background-reflectance as is the case

[5] Warren and Warren, *op. cit.*, 445-450.

with small stimulus-patches,[6] since surrounding the stimuli with black velvet had no appreciable effect on 'lightness' judgments.

Experiment IV shows that, when the contrast was asymmetrical (black velvet bordering on the gray strip but not on the white), an appreciable apparent lightening of the gray paper occurred. This apparent lightening is much greater for the dark gray paper against velvet (median judgment increased from 30 in Experiment I to 45) than for the light gray paper against velvet (median judgment increased from 65 to 70) as would be expected from the rules governing contrast. This experiment with asymmetrical contrast makes clear the fact that the use of large stimulus-fields reduces the effect of background-contrast to a negligible value only when homogeneous stimulus-surrounds are used.

The second judgments of *S*s made possible detection of misunderstanding of instructions. Since the second gray was either distinctly lighter or darker than the first, the second judgment should have indicated properly the direction of this difference if instructions were understood properly. According to this criterion, 75 out of 475 *S*s did not follow instructions, and their data were discarded.

If we compare the judgments obtained for a gray paper when it was presented second with the judgments obtained (for a different group) when it was presented first, a simple rule emerges. It appears that the second judgment was always shifted so as to increase the difference between the 'lightness' of the first and second grays.

Discussion

Subjective scales of gray have a long history. The early Munsell Value Scale of grays[7] (white = 10, black = 0) was so constructed that 'value' was equal to the square root of the percent of the incident light reflected by the gray.[8] Munsell may have been influenced by early work suggesting this relationship.[9] More recently, the report of a subcommittee appointed by the Optical Society of America indicated that the simple square-root relation between percent-reflectance and 'value' applied only when the gray stimulus-chips were viewed against a white background.[10] The modi-

---

[6] S. M. Newhall, D. Nickerson, and D. B. Judd, Final report of the O.S.A. subcommittee on the spacing of the Munsell Colors, *J. opt. Soc. Amer.*, 33, 1943, 385-418.

[7] A. H. Munsell, *The Atlas of the Munsell Color System*, 1915, Chart V.

[8] I. G. Priest, K. S. Gibson, and H. J. McNicholas, An examination of the Munsell Color System, *Tech. Papers. Bur. Standards*, No. 167, 1920.

[9] P. Breton, Mesure des sensations lumineuses, en fonction des quantités de luminière, *Comp. rend. Acad. Sci.*, 105, 1887, 426-429; A. Stefanini, Sulla legge di oscillazione dei diapason e sulla misura dell' intensità del suono, *Atti della R. Acc. Lucc. di Sc. Lett. ed Arti.*, 25, 1889, 305-400 (see especially pp. 383-391).

[10] Newhall, Nickerson, and Judd, *op. cit.*, 416.

fied Munsell Value Scale recommended by the subcommittee (and now in general use) deviates from the square-root relation since it was based upon viewing the stimuli against a gray background.

Thus, the recent literature indicates that there is no general relation between the reflectance of a gray and 'lightness' judgments, but only a series of subjective scales depending upon the background used for viewing the stimulus-papers. The mathematically simplest relation reported was for viewing against a white background, when 'lightness' was proportional to the square root of percent-reflectance.

The present experiment indicates that this square-root relation is not restricted to a single background, but is a general one when a large gray comparison-field is viewed next to a large white standard-field. Under these conditions the reflectance of a homogeneous surrounding field appears to have a negligible effect on 'lightness.' In the everyday lives of *Ss* they frequently make judgments involving the appearance of large fields contiguous in retinal image and at different distances from a light-source. Familiarity with these conditions may explain why 'lightness' judgments in the present experiment are in close quantitative agreement with the theory that the physical correlate of subjective visual intensity is the manner in which the luminance of an object changes with its distance from the light-source.

It is suggested that the 'value' judgments reported in the literature for small gray papers viewed against the background of another gray also have the same physical correlate as a basis, but these judgments are modified by contrast with the background and by the lack of a common border with a white stimulus. Such a border facilitates judgment by making possible a direct comparison of a gray relative to a white standard 'value.'

Recently, Stevens, Mack, and Stevens have commented upon the physical-correlate theory.[11] They state: "A 'theory' of this sort, since it is probably not susceptible to proof or disproof, must be judged, if at all, on the basis of the demands it makes upon our credulity." If by this statement these critics mean that the physical-correlate theory is not subject to experimental test, they are surely mistaken. Predictions based directly upon this theory have been tested and reported in the literature for brightness,[12] loudness,[13] and heaviness.[14] While none of these individual experiments furnished 'proof' of the general theory, each does provide evidence

[11] J. C. Stevens, J. D. Mack, and S. S. Stevens, Growth of sensation on seven continua as measured by force of handgrip, *J. exp. Psychol.*, 59, 1960, 60-67.

[12] Warren and Warren, *op. cit.*, 445-450.

[13] R. M. Warren, E. A. Sersen, and E. B. Pores, A basis for loudness-judgments, this JOURNAL, 71, 1958, 700-709.

[14] Warren and Warren, Effect of the relative volume of standard and comparison-object on half-heaviness judgments, this JOURNAL, 69, 1956, 640-643.

relating to its validity. The interested reader is referred to a summary of this and other pertinent evidence.[15]

Stevens, Mack, and Stevens state: "Another objection to this theory is that it is essentially negative. If taken seriously, it would seem to imply that neither magnitude estimations nor cross-modality matchings are able to tell us anything about the operating characteristics of a sensory system." The authors go on to say that, while it had been claimed that the physical-correlate theory provided a basis for Stevens' empirical law that power functions describe relations between stimulus and sensation-intensity, "the law itself would seem to lose much of its interest if it turned out that the characteristics of the sensory transducers have nothing to do with the exponents of the various power functions." Perhaps it is true that if the physical-correlate theory were valid, many of the psychological relations obtained through laborious experimentation would lose significance, but the over-all picture of psychophysics would be a much simpler one.

It seems important to emphasize here that the characteristics of the "sensory transducers" (which we take to mean receptors and their nerve fibers) *do not* follow the simple power-function relating sensation judgments to stimulus intensity. Direct electrophysiological measures do not appear to show this relation in any instance. In addition, changes in "sensory transducer" activity *are not* accompanied by corresponding changes in sensory intensity. Thus, in vision there is an abrupt change in sensory neural activity at the luminance corresponding to the transition from scotopic vision (involving only rods) to photopic vision (involving cones). This discontinuity is reflected by a change in the function relating just noticeable differences (*jnd*s), to stimulus-intensity,[16] yet there is no corresponding effect upon judgments of sensory intensity.[17] The central nervous system seems to transform extremely complicated sensory or input-relations concerning stimulus-intensity to a simple judgmental or output-relation. One wonders about Stevens, Mack, and Stevens' justification in considering that quantiative judgments concerning sensory-intensity may be used as a measure of the "characteristics of the sensory transducers." If it is their desire to investigate these characteristics, are there not valid techniques available for this purpose?

Conclusions

The present results are in agreement with predictions based upon the physical-correlate theory. The same relation of subjective judgments to the

[15] R. M. Warren, A basis for judgments of sensory intensity, this Journal, 71, 1958, 675-687.

[16] F. A. Geldard, *The Human Senses*, 1956, 30.

[17] R. M. Hanes, A scale of subjective brightness, *J. exp. Psychol.*, 39, 1949, 438-452; Warren and Warren, *op. cit.*, *J. opt. Soc. Amer.*, 48, 1958, 450.

square root of stimulus-intensity appears to apply for 'lightness' judgments with gray papers as for brightness-judgments with large contiguous fields. It is suggested that rather than consider 'bril' scales of subjective brightness and 'value' scales of subjective 'lightness,' or even combining these two into a single scale of subjective luminance, we simply consider these judgments as equivalent to estimates of a physical correlate, in this case the manner in which the luminance of an object changes with its distance from the source. This suggestion implies that judgments of sensory intensity are not related directly to the quantitative nature of the response of sensory nerves, but rather to central processes which correlate sensory input with previously observed physical relations of the outside world. We may consider this as another illustration of the statement by Helmholtz that "we are exceedingly well trained in finding out by our sensations the objective nature of the objects around us, but that we are completely unskilled in observing the sensations *per se*."[18]

SUMMARY

'Lightness' judgments of 475 *Ss* for various gray papers were obtained, either as numerical estimates relative to white or as marks placed on a line labelled 'white' at one end and 'black' at the other. The influence of 'prior stimulus-context' was avoided by considering only first judgments.

Subjective 'lightness' scales in current use are restricted to backgrounds of given reflectance, since under the viewing conditions employed, contrast with stimulus-surrounds has a considerable effect on judgments. The present study indicates that the influence of homogeneous surrounding fields on 'lightness' judgments is negligible when large contiguous fields are used.

It was found that, when the influence of background-contrast was eliminated, 'lightness' judgments (whether numerical 'ratio' judgments or linear 'interval' estimates) were proportional to the square root of the amount of light reflected by the gray paper. This relation corresponds to predictions based on the physical-correlate theory of sensory intensity. The physical correlate of 'lightness' appears to be distance; that is an individual's experience with the manner in which the distance of an object from its light-source determines the amount of light the object reflects. Equivalent predictions had been verified earlier for judgments of brightness of luminous fields, and it is suggested that subjective judgment of both brightness and 'lightness' have the same physical correlate.

Some conclusions relating to psychophysical scaling in general are discussed.

---

[18] H. v. Helmholtz, *Physiological Optics*, Vol. 3 (Translation edited by J. P. C. Southall), 1925, 9.

MEMORY IN MONKEYS FOR COMPOUND STIMULI

By LUCJAN S. STEPIEN and J. PIERRE CORDEAU, McGill University

The purpose of this paper is to describe the training of monkeys in a memory-task recently proposed by Konorski.[1] This new method requires that the animal discriminate between two compound stimuli, each of which is made up of two signals within the same modality and separated by a short interval. As used by Konorski, a compound stimulus, consisting of two tones identical in pitch, is positive and is reinforced by food if it elicits the appropriate response from the animal (placing a forelimb on the food-tray).[2] On the other hand, when the second signal is of a different pitch from that of the first, the stimulus is negative. It is not reinforced and the animal has to learn to leave the food-dispensing apparatus alone. In other words, the compounds S_x–S_x and S_y–S_y, etc., are positive, while S_x–S_y and S_y–S_x are negative as long as 'x' differs from 'y.'

The interesting feature in this procedure, and what distinguishes it from those previously used, is that here the first signal has no significance in itself; it can be part of either a positive or a negative compound stimulus. The correct response is possible only when the animal retains the trace of the first signal up to and including the time of the second signal, that comparison can be made between the two.

In the present series of experiments the monkeys were trained with both auditory and visual compound stimuli. The auditory stimuli were made up of clicks varying in frequency between 5 and 20 per sec.; two signals at similar frequencies constituted a positive stimulus, while signals with different frequencies denoted a negative stimulus. A similar arrangement held for visual stimuli using flashes of light instead of clicks. When training was completed, the animals underwent three types of tem-

* Received for publication July 27, 1959. This study was aided by a grant from the U.S.A. National Science Foundation. Dr. Stepien is a Fellow of the Rockefeller Foundation from the Nencki Institute of Experimental Biology, Warsaw, Poland, and Dr. Cordeau is a Medical Research Fellow of the National Research Council of Canada.

[1] Jerzy Konorski, A new method of physiological investigation of recent memory in animals, *Bull. Académie Polonaise des Sciences*, 7, 1959, 115-117.

[2] H. Chorazyna, Investigation of recent memory of acoustic stimuli in normal dogs, *idem*, 119-121.

poral lobe excisions. Postoperative behavioral and anatomical data are being reported separately.[3]

Method. The experiments were performed with naïve animals, five male and two female African Green monkeys (*cercopithecus athiops saboeus*). Their weights ranged from 2.5 to 5 kg. Each animal was trained and tested in its home cage. Once a day, this cage was brought into the experimental room where a wooden food-box, with a spring-hinged door, was attached to it. *E* sat behind a one-way vision screen and delivered the food-reward (pieces of apple and raisins) through a pipe leading into the food-box. Both auditory and visual stimuli were produced by a variable frequency generator connected to a loud speaker and a noiseless, neon-tube stroboscope. Both modes of stimulation could, therefore, be used separately or simultaneously. Once initiated, an automatic timer controlled the whole sequence of the compound stimulus. The delay between the two signals of each stimulus, as well as the duration of each signal, could be adjusted to any appropriate length of time. A manual two-way switch was also available for direct control of the stimulus-sequences. Throughout training no attempt was made to isolate the animal from the usual background noises of the laboratory.

Training procedure: (1) General. During the initial stages of training, the monkey was brought daily in its home cage to the experimental room for a period of approximately 1 hr. There it was shown successively: (1) that the food box contained food, the door being kept opened by means of a string; (2) that food could be obtained if the animal opened the door of the box and waited for food to be dropped through the pipe; (3) that food could only be obtained by opening the door after a buzzer had been sounded, door openings during the interval between buzzers not being reinforced. General habituation was considered as complete when, in a test-period of 50 successive buzzers at 15-sec. intervals, the animal produced 90% positive answers, with not more than 10 door openings during the intervals between buzzers. This criterion was reached after 10 to 15 days.

(2) Bell and buzzer discrimination. The next step in the training was to introduce a negative stimulus, a bell sound which was not reinforced by food and to which the monkey had to learn not to respond. Here again the animals were tested daily, each test-period consisting of 50 successive trials at intervals of 15 to 20 sec. The negative stimulus (bell) was presented after every third, fourth, or fifth positive stimulus (buzzer) and the test-period usually contained 40 positive and 10 negative stimulus-presentations.

It soon became apparent that the lack of reward (food) was not in itself sufficient punishment to teach the monkey not to open the door between trials or following the negative stimuli. Contrary to what had been observed by one of us with dogs undergoing similar training, the monkey does not seem to mind the absence of food when it opens the door on the negative signal, and appears to derive pleasure from the simple fact of playing with this interesting gadget.[4] The additional punishment introduced was a mild electric shock which the animals received when they touched the brass knob of the door during intertrial intervals or

[3] L. S. Stepien, J. P. Cordeau and T. Rasmussen, Effect of temporal lobe and hippocampal lesions on auditory and visual recent memory in monkeys, *Brain* (in press).

[4] H. Chorazyna and L. S. Stepien (in preparation).

on negative trials. Following this, the training proceeded smoothly and the animals were considered competent in this task when they had achieved 90% correct responses to the positive signals, with less than 10% errors on the negative signals and less than five movements during the intertrial for three consecutive test-periods. In the majority of cases the animals had learned by this time not to open the door during the interval between stimulus-presentations.

(3) Compound discrimination: (a) Auditory. As each animal reached criterion in the bell-buzzer discrimination, the training with compound auditory stimuli was begun. During the initial stages of training, the interval between the two signals of the same stimulus was short, *i.e.* approximately 1 sec. Fifty trials (with 10 negative stimuli inserted every fourth, fifth, or sixth trial) were given each day.

Initially, various frequencies for positive stimuli and two combinations S_8–S_{20} and S_{20}–S_8 for negative stimuli were used. After some weeks it became apparent that the problem was too difficult and could not be solved by the animals, or else that its solution would require an inordinately long time. The task was therefore simplified to two positive compounds S_8–S_8 and S_{20}–S_{20} and one negative stimulus S_{20}–S_8 (high-low sequence).

In this instance, also, punishment in the form of a mild electric shock followed incorrect responses in all test-trials, including door openings during or immediately after the presentation of the first signal of each compound stimulus, *i.e.* before presentation of the second signal. It became increasingly clear that the necessity for the animal to hold back his response for some 4–5 sec. after the beginning of each stimulus-sequence was one of the greatest obstacles to rapid learning encountered in the present study. Training proceeded on this basis and criterion was fixed at 95% correct responses to positive stimuli, with less than 5% errors on negative stimuli and less than three door openings during intertrial intervals for three consecutive test-periods.

After reaching criterion in this simplified version of the auditory discrimination, training was continued and the task made gradually more complex by the introduction for both positive and negative stimuli of the intermediate frequencies which had been eliminated at the beginning of the training period. At the same time, the reverse sequence of 'low-high' was reintroduced for negative stimuli and the intersignal interval of each stimulus gradually lengthened to 2, 3, and finally to 5 sec. A sample test-protocol for this later stage of training is shown in Table I. In all cases this gradual complication of the auditory task was carried out concurrently with training in visual discrimination as described below.

(b) Visual. The training in visual discrimination involved the same type of compound stimulus-conditions and similar frequencies to those used in the auditory sphere. For reasons stated below, two different approaches were used.

With four monkeys (SC_1, SC_2, SC_6, and SC_7), since considerable difficulties had been encountered in the initial training with compound auditory stimuli, it was thought advisable to introduce the intermittent photic stimulus during the later stages of auditory training so as to facilitate transfer from one sensory modality to the other. Therefore, to the clicks of the compound auditory stimulus we added the flash, flickering at the same frequencies and simultaneously with the auditory signals. After criterion had been reached with the two signals combined, they were then presented separately. The test-period now consisted of two successive series of trials each made up of 25 compound stimuli; first auditory, then visual.

With three monkeys (SC_3, SC_4, and SC_5), the visual stimuli were not presented with the auditory. When these animals had reached criterion in their auditory task, they were then given the visual stimuli. As in the previous case, each test-period now consisted of two successive series of trials each made up of 25 compound stimuli; first auditory, then visual.

In both groups, the visual task was then gradually made more complicated as in the case of the auditory, with the gradual introduction of more varied frequencies

TABLE I

SAMPLE TEST-PROTOCOL FOR SC_3

(Auditory discrimination, compound stimuli; signal duration, 2 sec.; interval between signals, 5 sec.; Test-period No. 70)

Trial	Time	Click frequencies	Response	Reinforcement
1	7:50	5— 5	+	+
2	7:51	10—10	+	+
3	7:52	20—20	+	+
4	7:53	20—20	+	+
5	7:54	20— 5*	0†	0
6	7:55	5— 5	+	+
7	7:56	20—20	+	+
8	7:57	20—10*	0	0
9	7:58	10—10	+	+
10	7:59	5— 5	+	+
11	7:60	20—20	+	+
12	8:01	10— 5*	0	0
13	8:02	5— 5	+	+
14	8:03	10—10	+	+
15	8:04	20—20	+	+
16	8:05	5—10*	0	0
17	8:06	10—10	+	+
18	8:07	20—20	+	+
19	8:08	20—20	+	+
20	8:09	10—10	+	+
21	8:10	10—20*	0	0
22	8:11	5— 5	+	+
23	8:12	10—10	+	+
24	8:13	20— 5*	0	0
25	8:14	5— 5	+	+

* Inhibitory stimulus-conditions. † 0=no response; correct.

for positive and negative stimuli, and with the gradual lengthening of the time-interval between the two signals of each compound stimulus.

Results. A sample curve of the performance of SC_4 in both auditory and visual discrimination is shown in Fig. 1. As stated above, the training of this animal was first begun on auditory discrimination (clicks) using various frequencies for positive compound stimuli (similar signals in the 5 to 20 per sec. frequency range), and two negative stimuli S_5–S_{20} and S_{20}–S_5 with an interval of 1 sec. between the two signals of each stimulus. After 20 test-periods, we realized that the problem was either too difficult for this animal to solve or that its solution would re-

quire an inordinately long time. The task was, therefore, simplified to two positive compound stimuli, S_5–S_5 and S_{20}–S_{20}, and one negative or inhibitory stimulus S_{20}–S_5 ('high-low'). It took 37 test-periods longer for this animal to reach criterion on this simplified version of the test. From then on the task was gradually made more complex, first by the introduction of the frequencies previously eliminated; secondly, by presentation of the reverse sequence 'low-high' for negative stimuli; and

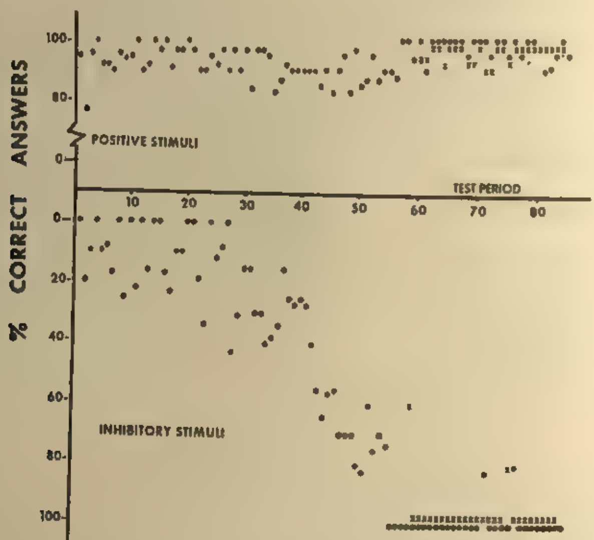


FIG. 1. PERFORMANCE OF SC_4 IN AUDITORY AND VISUAL DISCRIMINATION
• = auditory discrimination; × = visual discrimination

finally by a lengthening of the intersignal interval to 5 sec. It should be noted that once the animal had learned the simplified version of the test, the gradual increase in complexity did not alter its performance-level. In other words, it was able to generalize its previously acquired knowledge that similar signals meant a positive stimulus, while dissimilar signals made up an inhibitory stimulus.

Concurrently with this later stage of training in the auditory task, SC_4 was started independently on the visual task, using flashes of light instead of clicks, the signals remaining otherwise the same in time and frequency. Even though this animal had had no previous contact with intermittent photic stimulation, it transferred from one sensory modality to the other almost immediately. As can be seen in Fig. 1, SC_4, reached a

level of 60% correct answers to the inhibitory stimuli at the first test-period, and 100% thereafter.

Table II summarizes the performance of the seven monkeys trained in this manner, from the beginning of the experiment to the time they reached criterion in bell-buzzer discrimination and auditory compound-stimuli discrimination. In brackets under SC_3, SC_4, SC_5, and SC_6 are shown the number of test-periods during the initial stages of auditory training where the positive stimuli were made up of all frequencies in the 5 to 20 per sec. range, i.e. before the task was simplified to two positive stimuli and one negative stimulus as described above.

The results obtained in visual discrimination were identical in all ani-

TABLE II

NUMBER OF TEST-PERIODS AND TRIALS FROM BEGINNING OF TRAINING TO CRITERION

Discrimination	Animals						
	SC_1	SC_2	SC_3	SC_4	SC_5	SC_6	SC_7
Bell—Buzzer:							
(a) Total number of test-periods	29	26	15	39	12	15	46
(b) Positive stimulus-conditions	1277	1061	543	1464	447	543	1838
(c) Inhibitory stimulus-conditions	450	383	233	398	169	233	727
Auditory:							
(a) Total number of test-periods	36	33	57 (18)*	55 (20)	74 (22)	76 (37)	35
(b) Positive stimulus-conditions	1439	1285	2211	2201	2518	3634	1468
(c) Inhibitory stimulus-conditions	347	267	652	571	647	841	398
Visual:	Immediate transfer from auditory to visual discrimination except for SC_3						

* Figures in parentheses are the number of test-periods before the procedure was simplified.

mals except one (SC_3), regardless of the training method used. In the first group (SC_1, SC_2, SC_6, and SC_7), where criterion had been reached through simultaneous presentation of clicks and flashes, the separation of the two modes of stimulation did not affect correct performance with auditory cues. On the other hand, when the first few visual stimuli were presented alone, the animal was at a loss to know what to do. Transfer was quite easily obtained, however, and usually, by the end of the first test-period, every animal had become as proficient with visual as with auditory signals. The same can be said for two animals (SC_4 and SC_5) of the second group which were given the photic stimulus alone after they had reached criterion with the auditory signals. Both of these animals were able to perform their task with visual cues at the end of the first test-period. In one instance it was quite impossible to train the animal to perform with the flashing light. This monkey (SC_3) was afraid of the light; it became extremely nervous and refused to open the door to obtain its food following positive stimuli. No amount of habituation or

starving before test-periods could induce it to perform, hence training in visual discrimination with it was abandoned.

Discussion. The total number of test-periods required for the animals to reach criterion in the auditory discrimination of compound stimuli (Table II) should not be taken as a minimum but rather as a maximum. This is especially true of SC_3, SC_4, SC_5, and SC_6, who were started with positive compound stimuli of various frequencies for several test-periods before the task was simplified to the two positive stimuli S_{20}–S_{20} and S_5–S_5, and one negative stimulus S_{20}–S_5. Other factors were probably involved and contributed to the lengthy training period necessary before criterion could be reached. We think that it might be useful to list some of these.

In the first place, we believe that the task would have been easier if the monkeys had been offered a choice of two food-boxes rather than the 'go, no-go' situation used here. With stimuli made up of identical signals the animal would receive its food-reward from one food-box, while with stimuli of dissimilar signals the other food-box would be used. Secondly, since one of the great difficulties encountered here was to get the monkey to wait for the second signal before opening the door to the food-box, performance-levels might improve considerably if the animals were denied access to the feeding apparatus until the end of the second signal. For this purpose a sliding door, such as is described for many types of testing apparatus, could probably have been used to good advantage. Thirdly, the task might be easier if tones of different pitch had been used instead of click frequencies and colored lights instead of flash frequencies. The use of tones or colors instead of auditory or visual rhythms would have the additional advantage that the interval between the two signals of each stimulus could be made as small as desired, or the two could even be made to overlap during the initial stages of training or for postoperative testing. Since, as has been suggested, the monkey is a 'visual' rather than an 'auditory' animal, the reverse sequence of training, that is, visual discrimination before auditory, might also have facilitated learning and have speeded up the time necessary to reach criterion. If, however, colors and tones were used instead of rhythmic frequencies, one could not expect the immediate transfer of knowledge from one sensory modality to the other which has been reported here.

Similar experiments were performed in Konorski's laboratory by Chorazyna, who used dogs as experimental animals and tones of similar or different pitch as posi-

tive and negative stimuli. It is worthwhile to note that a problem of this kind appears to be more easily solved by monkeys. Whereas the dog seems to be incapable of generalizing from a 'low-high' pitch-sequence as negative stimulus to one of 'high-low' pitch, the monkey readily transfers from one to another dissimilar stimulus no matter what specific frequencies happen to be involved.

The phenomena of transfer from one sensory modality to another, from the auditory to the visual sphere in the present circumstances, deserves brief mention. As reported above, after training in auditory discrimination the animals were switched to visual discrimination with signals similar in infrequency to the clicks. We thought that this new training would be shorter than, and facilitated by, the previous experience of these animals in auditory compound stimulation discrimination, but we were not prepared for the extreme rapidity with which this transfer occurred. In all animals but one it took less than one training session made up of 20 positive and 5 negative stimuli for the differentiation to be complete in the visual field. This should be compared to the more than a thousand positive and several hundred negative stimuli which were necessary to obtain differentiation in the auditory training. The rapidity of this phenomenon can only be accounted for by an active physiological process which facilitates the transfer of memory traces of rhythms from one sensory modality to another. In a discussion on this topic, Konorski has speculated that such transfer phenomena involve structures at a 'higher level' of nervous activity than the specific sensory projection-areas, whereby the rhythm is so to speak abstracted by the animal from the other characteristics of a given stimulus.[5] In view of the recent electrophysiological findings in conditioning, the possibility that subcortical centers play a major role in transfer of this type should be kept in mind and made the subject of further investigation.

[5] Konorski, *Conditioned Reflexes and Neuron Organization*, 1948, 161 f.

A TEST OF SPENCE'S THEORY OF INCENTIVE-MOTIVATION

By R. C. GONZALEZ, Bryn Mawr College, and LEONARD DIAMOND
University of Maryland

The abrupt shifts in the running-speed of rats, found by Crespi to accompany sudden variations in amount of incentive,[1] led Hull, in the latter versions of his system,[2] to abandon the hypothesis that habit-strength varies with magnitude of reward. Instead, Hull conceived of amount of reward as affecting the strength of an incentive-motivational factor, *K*, which, like drive, was assumed to interact multiplicatively with habit-strength in the determination of reaction-potential. He did not attempt to specify the mechanism underlying the action of *K* beyond stating that it varied directly with amount of reward.

Hull's analysis has been extended recently by Spence, who suggests the fractional anticipatory goal-response as the basis of *K*.[3] Through generalization, Spence assumes, any stimuli resembling those to which a consumatory response has been classically conditioned come to evoke noncompetitional fractions of the goal-reaction. The occurrence of these fractional goal-responses in the absence of food is assumed to contribute to the over-all level of motivation, and the value of *K* is determined by the vigor with which these responses are made.

Assumed, as it is, to depend upon classical conditioning and in principle, therefore, to be capable of manipulation independently of instrumental learning, *K* seemingly provides the basis for deducing such phenomena as those of latent learning. Indeed, following a suggestion made by Spence,[4] Stein has tested his theory of incentive-motivation in the context of what is essentially a latent-learning experiment.[5] Two groups of rats were given a

* Received for publication September 16, 1959. This experiment was conducted in the Laboratory of Psychopharmacology at the University of Maryland.

[1] L. P. Crespi, Quantitative variation of incentive and performance in the white rat, this JOURNAL, 55, 1942, 467-517.

[2] C. L. Hull, Behavior postulates and corollaries, *Psychol. Rev.*, 57, 1950, 173-180; *Essentials of Behavior*, 1951, 47-51; *A Behavior System*, 1952, 140-148.

[3] K. W. Spence, Theoretical interpretations of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 721-722; The relation of response latency and speed to the intervening variables and N in S-R theory, *Psychol. Rev.*, 61, 1954, 209-215; *Behavior Theory and Conditioning*, 1956, 134-164.

[4] Spence, *Behavior Theory and Conditioning*, 147-148.

[5] Larry Stein, The classical conditioning of the consummatory response as a determinant of instrumental performance, *J. comp. physiol. Psychol.*, 50, 1957, 269-278.

series of unrewarded trials in a runway, *Rx*, which was similar to the empty end-box, *Bx*. Then one group was fed in *Bx* and the other in *By*, a goal-box quite unlike *Rx* and *Bx*. Sharp increases in speed of running following the direct feedings were anticipated on the assumption that the r_g conditioned to the goal-box stimuli would generalize to the runway and thus increase the value of K. It was predicted that the speed of running of *Ss* fed in *Bx* would, because of greater generalization, exceed that of *Ss* fed in *By*, and rapidly approach the level of a control group given rewarded runs to *Bx* from the very beginning.

Of course, a cognitive theory such as Tolman's leads to precisely the same predictions.[6] With the expectation $\dot{S}_{Rx}\dot{R}\rightarrow\dot{S}_{Bx}$ established in the first stage of the experiment, the increase in the valence of $\dot{S}_{Bx}$ produced in the second stage by the feeding in *Bx* should be manifested in the third stage by an increase in speed of running. From Tolman's point of view, the magnitude of the increase in speed must vary directly with the similarity of the goal-boxes encountered in the first two stages of training; from Spence's point of view, it is the similarity of the goal-box of the second stage and the runway which is critical. The two points of view lead to the same predictions in Stein's experiment because the goal-box encountered in the second stage was either similar or dissimilar *both* to the runway and to the goal-box of the first stage. With two such divergent theories suggesting the same outcome, the results of Stein's experiment are especially interesting: neither group showed a significant increase in speed of running.

The experiment to be reported here, which is an extension of Stein's, was intended to clarify this striking failure of prediction. For one group of rats, the runway and goal-box of the first stage were similar (the *RxBx* situation), while for a second group they were dissimilar (the *RxBy* situation). In the second stage, half the *Ss* in each group were fed in *Bx* and the remainder were fed in *By*, yielding four subgroups: (a) *RxBx-Bx*, (b) *RxBy-Bx*, (c) *RxBy-By*, and (d) *RxBx-By*. Groups *a* and *d* are equivalent to Stein's, while groups *b* and *c* lead Spence and Tolman to *opposed* predictions. According to Spence, the animals of group *b* should run faster in the third stage, but, according to Tolman, the animals of group *c* should run faster.

This experiment, which was suggested recently by Bitterman,[7] yields a meaningful pattern of results in the light of which the two theories may be seen in clearer perspective.

---

[6] E. C. Tolman, Principles of purposive behavior, in Sigmund Koch (ed.), *Psychology: A Study of Science*, 2, 1959, 92-152.

[7] M. E. Bitterman, Review of Spence's *Behavior Theory and Conditioning*, this Journal, 70, 1957, 144.

Method

Subjects. The *Ss* were 32 experimentally naïve, male, hooded rats of the Royal Victoria strain, about 120 days old at the beginning of the experiment.

Apparatus. The apparatus employed consisted of a starting-box, an enclosed runway, and two interchangeable goal-boxes which could be placed at right-angles to either side of the distal end of the runway. (The right-angle turn made the interior of the goal-box invisible from the runway.) The runway was 4 in. wide, 48 in. long, and 6 in. high. The starting-box was $6 \times 9 \times 6$ in., and the goal-box, $6 \times 12 \times 6$ in. The runway had a wire-mesh cover, while the starting-box and goal-boxes were covered with Plexiglas. The wall of the starting-box adjoining the runway was a sliding door which was operated manually by *E*. A second sliding door was set into the runway 4 in. before the goal-compartment. Starting-latencies and running-times were recorded on separate Standard Electric timers which were graduated in units of 0.01 sec. Opening the starting-box door activated a microswitch which started the first clock. Closing this door released the microswitch, stopping the first clock and activating the second. Closing the second door stopped the second clock. At the beginning of each trial, the starting-box door was closed, and the runway door was open. Two sets of apparatus were employed. The walls and floors of one set were painted a flat black, and the walls and floors of the second were painted a flat white. In all other respects the two sets were identical.

Procedure. The experiment was conducted in two phases. Prior to preliminary training, the 32 *Ss* were randomly assigned to four groups, each of which contained 8 animals. First, 16 *Ss*, 4 in each group (again selected randomly), were carried through to the completion of the experiment. The experiment then was replicated with the remaining 16 *Ss*.

Preliminary training. The *Ss* were placed on a 24-hr. feeding schedule and were taken once daily, for 15 days, into the experimental room for 20-min. periods to adjust to handling. Each *S* was fed two 6-oz. pellets of Purina lab chow in its home-cage 2 hr. after each day's handling.

Stage 1. Following the preliminary-training period, there were 20 training trials, 4 on each of 5 successive days. Each day *S* was brought into the experimental room and allowed to adapt for 2 min. in the carrying cage. On each trial, *S* was placed in the starting-box and, when it oriented toward the door, the door was opened. The door was closed again as soon as *S* entered the runway. In the same manner, the door at the end of the runway, which was open at the beginning of each trial, was closed behind the animal. *S* remained in the goal-box for 1 min., after which it was returned to the starting-box for the beginning of the next trial. The goal-box was placed on the right side of the runway (that is, entered by a right turn) on half of each day's trials and on the left side (entered by a left turn) on the other half, the order of placement being varied systematically in accordance with prearranged orders. All training trials were unrewarded. At the completion of each day's trials, *S* was returned to its home-cage where it was fed two 6-oz. pellets 2 hr. later.

Stage 2. During the second stage of training, each *S* was placed in one of the goal-boxes and fed, four times on the day following the last training day, and once on the second. *S* was brought into the experimental room, allowed 2 min. to adapt in the carrying cage, and then placed directly into the goal-box containing a small cup of wet mash. A partition, painted the same color as the goal-box, prevented entrance into the alley. After 4 min., *S* was removed from the goal-box and placed

in the carrying cage for 1 min., after which it was placed again in the goal-box for 4 min., and so forth. The food thus consumed on the first four feedings was approximately equal to *S*'s daily ration. No food was given in the home-cages during this phase of the experiment. The fifth direct feeding was given in the same manner on the following day.

The experimental conditions to which each group was assigned are summarized in Table I. Group A was run in *RxBx* and fed in *Bx;* Group B was run in *RxBy* and fed in *Bx;* Group C was run in *RxBy* and fed in *By;* Group D was run in *RxBx* and fed in *By*. For half of each group, *x* was white and *y* black; while for the remaining *S*s, *x* was black and *y* was white.

Stage 3. Immediately following the fifth direct feeding, *S* was removed from the goal-box and taken to its home-cage. Returned to the experimental room 15 min. later, *S* was allowed 2 min. in which to adapt, and then placed in the starting-box

TABLE I

DESIGN OF THE EXPERIMENT

Group	N	Runway, stage 1	Goal-box, stage 1	Goal-box, stage 2
A. *RxBx–Bx*	4	black	black	black
	4	white	white	white
B. *RxBy–Bx*	4	black	white	black
	4	white	black	white
C. *RxBy–By*	4	black	white	white
	4	white	black	black
D. *RxBx–By*	4	black	black	white
	4	white	white	black

for the first test-trial. The apparatus and procedure for each *S* were exactly the same as in training. In all, eight trials were given each *S* in the third stage, four on each of two successive days.

RESULTS

The performance of the four groups during the first and third stages of the experiment is shown in Fig. 1 in terms of mean log running time for blocks of four trials. The values for the first trial of Stage 3 are so close to those for the first day that a trial-by-trial plot is not presented. The starting-latencies, which are not shown, show the same relative pattern as do the running times.

The four groups did not differ significantly in running-times during the training period ($F = 0.76$, $p > 0.05$). The effects of the direct-feedings on subsequent running-times were evaluated with *t*-tests of the difference between final level of performance during Stage 1 and initial level during Stage 3. For each group, the final trial and the last block of four training trials in Stage 1 were compared respectively with the first trial and with the first block of four trials in Stage 3. The decreases in running times were significant in Groups C ($t = 3.1$ $p < 0.05$) and D ($t = 2.95$, $p < 0.05$). The decrease in Group A approached, but did not reach, significance ($t =$

1.95). The *Ss* in Group B showed a slight *increase* in running time. The groups also differed in their pattern of running times over the eight trials of Stage 3. The times for Groups A and C tended to decrease, while those for Groups B and D tended to increase. These results put the theoretical differences between Tolman and Spence into clear perspective.

From the standpoint of Spence's theory, the performance during Stage 3 of Groups A and B, the groups which were fed in a goal-box similar to the

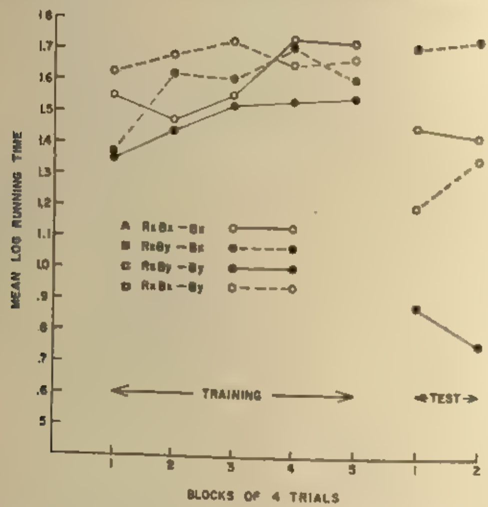


FIG. 1. PERFORMANCES DURING STAGES 1 AND 3

runway, should have been superior to that of Groups C and D, which were fed in a goal-box dissimilar to the runway. A glance at Fig. 1 reveals that the results are in precisely the opposite direction. The inadequacy of Spence's theory is most strikingly revealed by a comparison of the performance of Groups B and C, with respect to which his theory contrasts most sharply with the cognitive view. Group B was fed in a goal-box similar to the runway but unlike the goal-box originally encountered. Group C was fed in a goal-box dissimilar to the runway but similar to the goal-box originally encountered. According to Spence, generalization of r_g to the runway-stimuli should be greater in Group B, the value of *K* should therefore be larger, and the decrease in running time should be greater for *Ss* in this group. Tolman's theory suggests, however, that the *Ss* in Group C, having associated food with the goal-box previously encountered at the end of the runway, should be superior. The results are consistent with the cognitive interpretation.

Although the results of this experiment point convincingly to the correctness of Tolman's emphasis, they suggest the incompleteness of his

analysis. Speed of running in Stage 3 was found to vary directly with the similarity of the goal-boxes in the first two stages of training and inversely with the similarity of the goal-box and the runway in Stage 2. Tolman's theory stresses the importance of the first relationship and the results are, in this respect, consistent with this emphasis. It does not, however, make explicit allowance for the significance of the second relationship and, to this extent, the theory is lacking. Unlike the inadequacy of Spence's theory, however, which seems rather central, the deficiency in Tolman's theory may be remedied by a simple extension which not only brings the theory into line with the data but which adds to its deductive power as well.

From the standpoint of Tolman's theory, the contiguous occurrence of S_1-R_1-S_2 leads to the establishment of the expectation $S_1\dot{R}_1 \rightarrow \dot{S}_2$. If then S_2, or a similar stimulus, S'_2, is paired with food, S_2 acquires positive valence. The valence of S_2 increases with its similarity to S'_2, and subsequent speed of running, in turn, varies directly with the valence of S_2. Of course, stimuli other than S_2, to the extent that they are similar to S'_2, also acquire positive valence. Clearly, this formulation may be extended logically to include an account of the effects on the properties of S_1 of the feedings in S'_2. To the extent that S_1 resembles S'_2, it also must acquire some positive valence. Viewed in these terms, subsequent speed of running is seen to be a function, not of the absolute valence of S_2, but of its valence relative to the valence of S_1.

Interpreted in this manner, the results of all the groups in the present experiment become understandable. In Group C, S'_2 is much more similar to S_2 than to S_1, the goal-box valence accordingly is great relative to the runway-valence, and *S*s show a marked decrement in running time. In Group A, the goal-box valence is the same as in Group C, but, because S_1 is more similar to S'_2 in this group, the *relative* magnitude of the goal-box valence is reduced and the *S*s, therefore, show only a moderate decrement in running time. The *S*s of Group B, by contrast, show an *increase* in running time because the *runway* valence exceeds that of the goal-box. The seemingly paradoxical results for Group D yield equally well to this analysis. Although S'_2 is quite distinct both from S_2 and from S_1 with respect to color, its structural properties are quite similar to those of S_2. The valence acquired by S_2, therefore, although moderate, is substantial in relation to that acquired by S_1, and the *S*s of this group show a significant decrease in running time.

Table II is presented to illustrate the essential features of this interpretation. Strength of response on the first trial of Stage 3 is assumed to be a function of the ratio of the goal-box valence to the runway-valence. The valence of each of these stimulus-complexes is specified in terms of its similarity to S'_2 and similarity, in turn, is specified in terms of the two chief dimensions with respect to which the stimuli may be compared—color and

structure. In the assignment of numerical weights, the brightness-dimension is more heavily emphasized on the assumption that black and white, as extremes on the dimension, are perceptually more distinct than are the structural properties of the components of the apparatus (which could be varied to a much greater extent). The values assigned are admittedly arbitrary, as is the manner in which the two dimensions are assumed to interact in determining similarity. Precise specification is, however, open to empirical determination, and the purpose of assigning values, in any event, is merely to illustrate the extension of Tolman's theory as it applies to the results of the present experiment.

With Tolman's theory thus extended and shown to be capable of accounting for the results, the implications of the present findings for Spence's

TABLE II

THEORETICAL STRENGTH OF RESPONSE IN STAGE 3

(A similarity-value of 1.0 indicates identity of stimuli with respect to the dimension in question. The similarity of black and white is assigned a value of 0.1 and the similarity of structure of goal-box and runway is assigned a value of 0.4. The sums shown in the seventh and tenth columns represent the valences of goal-box and runway, respectively. Decrement in running time is based on the differences between the mean of block 5 in stage 1 and the mean of block 1 in stage 3.)

Group	S_1	S_2	S_2'	S_2–S_2' Similarity		Sum	S_1–S_2' Similarity		Sum	Ratio	Running-time decrement
				color	structure		color	structure			
A.	*Rx*	*Bx*	*Bx*	1.0	1.0	2.0	1.0	0.4	1.4	1.43	0.25
B.	*Rx*	*By*	*Bx*	0.1	1.0	1.1	1.0	0.4	1.4	0.79	−0.08
C.	*Rx*	*By*	*By*	1.0	1.0	2.0	0.1	0.4	0.5	4.00	0.66
D.	*Rx*	*Bx*	*By*	0.1	1.0	1.1	0.1	0.4	0.5	2.20	0.49

theory may be more fully considered. From the point of view of Spence's theory, the properties of S_2 are represented by implicit goal-responses which are evoked, through stimulus-generalization, by stimuli which resemble those to which the goal-responses have been conditioned. The theory, therefore, requires a positive relationship between speed of running and the similarity of S_1 to S'_2. To account for the negative relationship obtained in the present experiment, it is, of course, possible to invoke the notion of competing responses—it may be assumed that r_g somehow conflicts with running—but then an interpretation of K in terms of r_g must be abandoned. As Spence saw clearly, r_g, if it is to be the basis of K, must be assumed not to conflict with instrumental activity.[8] It is interesting, therefore, to

[8] Nevertheless, with only his own results before him, Stein (*op. cit.* 269-278) attempted to account for them in terms of a modification of Spence's theory based on competing responses. He proposed that r_g, having occurred in the runway for the first time in Stage 3, may have effectively competed with running. On the assumption that, with repeated evocations in the runway, prior to testing, the competitive effects of r_g would dissipate, thus permitting its effect on K to be manifested, Stein looked for increased speed of running in a second experiment in which *Ss* were given directplacement feedings in *Bx* 15 min. after each unrewarded training-trial. The results of this experiment also were negative.

speculate on how proponents of S—R theory might account for the present results. One possibility is to reject Spence's conception of the mechanism underlying the action of K in favor of Hull's version, and to invoke the notion of competing responses. The value of K then would be assumed to be equal for all groups, and speed of running would vary inversely with the magnitude of the competing responses. This interpretation accounts for the superiority of Groups C and D, although it requires that their performance be equal (as should that of Groups A and B). It is important to note, however, that, even in these terms, no account of the dependence of performance on the relationship of S_2 and S'_2 is possible. The similarity of such a formulation to expectancy theory is, therefore, only superficial. It could not, for example, provide an adequate account of behavior in situations in which the relationship of S'_2 to S_1 is held constant while that of S'_2 to S_2 is varied.

Whatever S—R interpretation may be given to these results, they clearly will not yield to analysis in terms of Spence's theory of incentive-motivation. They require a radical revision of his conception of the mechanism underlying the action of reward, and cast serious doubt on the validity of interpretations of other phenomena which are based on this concept. Phenomena of the type here investigated are readily understood in cognitive terms, and any serious attempt to account for them in S—R terms must await a more plausible theory than the one advanced by Spence.

Summary

An experiment is reported which clarifies the ambiguous results of a previous test of Spence's theory of incentive-motivation. Four groups of eight rats each were given a number of unreinforced trials in a runway leading to an empty goal-box. The runway and goal-box were similar for two groups and dissimilar for the remaining two. In the second stage of the experiment, each S was placed directly into a box and fed. From the standpoint of Spence's theory, the effectiveness of the directfeedings (measured by subsequent increases in speed of running) was expected to vary with similarity of the goal-box of the second stage to the runway. From Tolman's cognitive viewpoint, change in running speed was expected to vary with the similarity of the goal-boxes encountered in the first two stages of training.

The results, which directly contradicted Spence's theory, pointed to the correctness of Tolman's emphasis, but suggested the incompleteness of his analysis. An extension of Tolman's theory was proposed to account for the results, and the implications of the findings for S—R theory were discussed.

JUDGMENTS OF VELOCITY AND WEIGHT IN A CAUSAL SITUATION

By THOMAS NATSOULAS, Wesleyan University

Michotte has conduced a number of experiments on perceptual phenomena which he calls launching and releasing (*lancement* and *déclenchement*).[1] These studies primarily involved a simple stimulus-situation consisting of the apparent head-on collision of two differently colored squares, one of which (*A*) is stationary after the collision, while the other square (*B*) is motionless before *A* collides with it.[2] If the period of *A*'s contact with *B* is within certain limits, either launching or releasing is reported by *S*, depending on the relative velocities of the two squares. When *B*'s velocity was 1.8 or more times *A*'s velocity, three practiced *S*s reported, some of the time, that upon contact *A* seemed to release or trigger off a force in *B* which was responsible for *B*'s movement. When *B*'s velocity was 1.6 or fewer times *A*'s velocity, the same *S*s always reported that *A* appeared to produce *B*'s movement directly, that the force of *A*'s collision with *B* was responsible for *B*'s movement. The former report Michotte calls releasing, the latter, launching.

The theory proposed by Michotte to account for the launching and releasing phenomena has as its goal the description of the perceptual experiences which lead his *S*s to make their reports.[3] His results, on the other hand, are related almost exclusively to physical variables rather than to the *S*s' judgments of the physical variables. An important factor in the theory of 'launching-releasing' is the relative velocities of *A* and *B*. In view of this it may prove fruitful to examine the velocity ratios from a psychophysical point of view.

As Michotte points out, furthermore, there are exceptions to these results.[4] Some *S*s report releasing when the velocities of the two squares are equal. Michotte's attempt to account for these exceptions in terms of apparent velocities suggests that a closer relationship may exist between

* Received for publication October 1, 1959. This article is based on a dissertation submitted to the Department of Psychology of the University of Michigan in partial fulfillment of the requirements for the Ph.D. degree. The author is indebted to Dr. J. David Birch for his advice and assistance.

[1] Alberto Michotte, *La perception de la causalité*, 2nd ed., 1954, 1-306.

[2] Michotte, *op. cit.*, 106 ff.

[3] Michotte, *ibid.*, Chap. VIII.

[4] Michotte, *ibid.*, 107.

the reports and judged velocity-ratios than Michotte found between the reports and physically defined velocity-ratios.

In the present experiment both the physical velocity and size-ratios of *A* to *B* are varied. The *S*s are required to estimate either the velocity-ratios or the weight-ratios of *A* to *B*. Estimates of velocity-ratios are collected for the reasons stated above. Estimates of weight-ratios are collected with the expectation that findings in Michotte's typical launching-releasing situation may be ordered by psychological principles which correspond to those governing the physics of collisions. The latter principles include as variables the velocities and the masses of the colliding bodies. To insure variation in estimates of weight-ratio across stimulus-objects, the sizes of *A* and *B* are systematically varied.

Method and Procedure

Stimulus-objects. All the stimulus-objects used have the following aspects in common. *A*, a black square or rectangle, appears at the left end of an aperture, 4.5 cm. from its center. *B*, a red square or rectangle, appears at the center. Both are stationary for 0.27 sec. *A* then commences to move toward *B*. *A*, having traversed the 4.5 cm., comes into contact with stationary *B*. They remain in stationary contact for 0.03 sec. Then *B* crosses 4.5 cm. to the right, at which point the aperture ends and *B* disappears behind the screen. *A*, too, disappears at its position in the center of the aperture. This sequence of events occurs 10 times for each stimulus-condition on each trial, the 10 presentations taking place in 15 sec.

All combinations of 5 velocity-ratios and 5 size-ratios are used yielding a total of 25 different conditions. The velocity-ratios of *A* to *B* are 3:1, 2:1, 1:1, 1:2, and 1:3. Three velocities are used to produce these ratios: 60, 40, and 20 cm./sec. For example, a 3:1 velocity-ratio means that *A* moves to meet *B* at 60 cm./sec. and that *B* moves off after contact at 20 cm./sec. *A* and *B* vary with respect to size in the same five ratios. In all cases *A* and *B* are 0.5-cm. high (vertical). They may have one of three widths (longitudinal): 0.5, 1.0, or 1.5 cm. Thus, for example, a size-ratio of 1:2 indicates that *A* is a 0.5-cm. square while *B* is a rectangle 0.5-cm. high and 1.0-cm. wide.

Apparatus. The stimulus-effects described are produced by the use of disks.[5] Cardboard disks rotated individually behind a screen by a motor at 40 r.p.m., each produce one stimulus-situation. The screen, which stands 2.5 cm. in front of the rotating disk, has a horizontal aperture, 11-cm. wide and 0.5-cm. high. Through the aperture an oblong segment of the disk, illumined directly by a 100-w. lamp, can be seen. Overhead and daylight illumination further brighten the experimental cubicle.

On each of the disks two bands are painted, one red and one black. The three continuous segments which comprise the black band have shapes which may be described as follows: (a) a portion in the shape of an arc of the disk, extending over 65° of the disk, the band's inner border 23.5 cm. from the disk's center; (b) a curved portion covering 18°, 27°, or 54° (depending on the desired velocity of

[5] Michotte, *ibid.*, 26-32.

A), approaching uniformly the center of the disk, with the band's inner border ending up 19 cm. from the disk's center; and (c) another arc-shaped portion extending over 60.7° of the disk, with the band's inner border 19 cm. from the center of the disk. The black band is made up of these three parts in a continuous line in the above order. The result, when the black band is made to rotate behind the aperture at 40 r.p.m., is respectively (a) A's being stationary for 0.27 sec., (b) A's moving to the right at 60, 40, or 20 cm./sec. for 4.5 cm., and (c) A's remaining stationary for 0.25 sec.

Extending over the same segment of each disk as does the black band is a red band which is responsible for B and its movement. It consists of two continuous parts in the following order: (a) an arc-shaped portion beginning parallel with the black band and ending 6.7° beyond the point at which A enters its second stationary stage, with the band's outer border 19 cm. from the disk's center; and (b) a curved portion which extends over 18°, 27°, or 54° of the disk (depending on the desired velocity of B), approaching uniformly the center of the disk, with the band's outer border ending 14.5 cm. from the center. When the disk rotates behind the aperture at 40 r.p.m. the result is, respectively, (a) B's being stationary from 0.35 to 0.50 sec. (depending on the velocity of A), and (b) B's moving to the right at 60, 40, or 20 cm./sec. for 4.5 cm.

The third portion of the black band and the first portion of the red one border on each other; both the inner border of the black band and the outer one of the red band lie 19 cm. from the disk's center. This occurs over only 6.7° and produces a period of stationary contact of A with B of 0.03 sec.

With respect to the rotating disk, the position of the aperture in the screen is such as to expose a 0.5-cm. high area along a horizontal radius of the disk. The exposed area runs from a point 25.5 cm. from the center of the disk to 14.5 cm. from the center of the disk.

Subjects. Two groups of 19 men each—students in the elementary psychology at the University of Michigan—participated in the study. All served without knowledge of the problem. They were assigned arbitrarily to one of the two experimental groups.

Procedure. The *Ss* sat 5 ft. from the apparatus with the aperture at eye-level. They participated either individually or in pairs. When in pairs, they sat side by side with a cardboard screen separating them. They could not see each other nor each other's responses. Only three members in each group served individually.

Instructions. The following instructions were read to both groups:

This is a study in the perception of movement. You will see in the aperture two small squares or rectangles, A (a black one) and B (a red one). They will move across the aperture at various velocities and will seem to be of various weights. A will move from this end of the aperture to the center where it will meet B. B will then move from the center to the right and disappear behind the screen. They will then reappear at their original positions and then go again through the same movements as before.

The *Ss* of Group I [II] were then given the following specific instructions:

Your task will be to judge the relative velocities [weights] of A and B. On the sheet before you, enter for each trial the ratio of the two velocities [weights]. Decide which one is the faster [heavier] and indicate how many times faster [heavier] than the other object it seems to be. There will be 75 trials for you to judge, and

you will see each one 10 times. Watch carefully each trial; you will have plenty of time to write down your response while I change the stimulus-conditions. At the end of each set of 25 trials we will have a 5-min. break. Any questions?

Questions were answered whenever possible by repeating a part of the above instructions. The *Ss* of Group II, who were instructed to give judgments of weight,

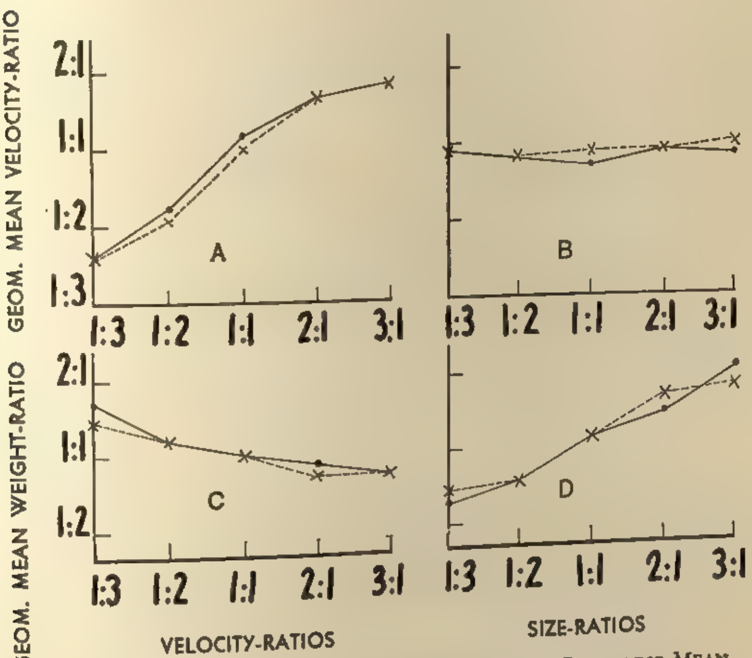


FIG. 1. THE RELATIONSHIPS BETWEEN JUDGMENTS OF GEOMETRIC MEAN VELOCITY-RATIOS AND WEIGHT-RATIOS AND PHYSICAL VARIABLES
(Trial 2, solid lines; Trial 3, broken lines)

asked more questions about their task than did those of Group I. It was consequently, necessary to add the following to the instructions for Group II:

We often judge weight not only by holding or touching things, but also visually. This is your task: judge the relative weights of *A* and *B* on the basis of what you see.

The stimulus-conditions were presented three times, each time in a different order. All three orders were counterbalanced with respect to the two stimulus-dimensions. Within each successive group of five conditions there being five such groups in each order, each velocity-ratio and each size-ratio occurred only once. Furthermore, no two stimulus-conditions which had the same velocity- or size-ratio were given successively. For the 10 presentations of each condition, 15 sec. were necessary. The removal of a disk and the substitution of another took 45 sec. Each series of 25 judgments, therefore, took 25 min.

Results

Trial 1 was treated as a practice trial and not submitted to statistical analysis. Trials 2 and 3 are analyzed separately, for each of the two experimental groups. The data are comprised of judgments of velocity-ratio from Group I and weight-ratio from Group II. Each of the four analyses is a three-way analysis of variance: velocity-ratios (*VR*), size-ratios (*SR*), and *S*s. The measures used for the analyses are logarithms to the base 10 of velocity- or weight-ratio, with 1.000 added to each entry in order to make all of them positive.

Velocity-ratio judgments. Fig. 1A shows the relationship between the velocity-ratios of *A*:*B* and the geometric mean velocity-ratio judgments for Trials 2 and 3. Each point in this figure is the geometric mean of 19

TABLE I
Analysis of Judgments: Variance of Velocity-Ratios
(Log$_{10}$ plus 1.000 transformation)

Source		Trial 2		Trial 3	
	df.	*MS*	*F**	*MS*	*F**
Velocity-ratios (*VR*)	4	6.8895	126.65§	6.8947	87.16§
Size-ratios (*SR*)	4	.0708	2.43	.0250	1.80
*S*s	18	.0450	1.08	.0514	1.10
VR×*SR*	16	.0381	2.03†	.0424	2.83‡
VR×*S*s	72	.0544	2.89§	.0791	5.27§
SR×*S*s	72	.0291	1.55‡	.0139	.93
Residual	288	.0188		.0150	

* For interactions, *MS* tested against residual *MS*. *MS*s of *VR* and *SR* tested respectively against *VR*×*S*s, and *SR*×*S*s. *MS* of *S*s tested against pooling of *VR*×*S*s and *SR*×*S*s.
† Significant at 5% level.
‡ Significant at 1% level.
§ Significant at $\frac{1}{10}$% level.

*S*s responding to five stimulus-conditions, all of which have an equal *A*:*B* velocity-ratio but differ in their ratios of sizes of *A*:*B*. The relationship is a direct one: as the velocity-ratios increase from 1:3 to 3:1 the ratio-judgments also increase. The ratios as judged, however, appear to be contracted in range.

Fig. 1B shows the relationships between the geometric mean judgments of velocity-ratios and size-ratios. The lines for both Trials 2 and 3 are very close to horizontal.

The two analyses of variance performed on the log velocity-ratios are summarized in Table I. In all except one case, the first-order interactions are statistically significant. What seems to be responsible for this state of affairs is that the residual *MS* is of such small magnitude that even slight variation in results across levels will show up as significant. Plots (not shown here) of the geometric means relevant to the *VR* × *SR* interaction show no systematic differences from level to level.

Only the velocity-ratios, among the main effects, seem to be a determinant of the log judgment of velocity-ratio when tested against the appropriate first-order interaction. The *F*-values for the main effects of both the size-ratios and *S*s are not significant.

Weight-ratio judgments. Fig. 1D shows that the relationship between the size-ratios and geometric mean weight-ratio judgments is a positive one. On the other hand, as can be seen in Fig. 1C, the relationship of the velocity-ratios to the geometric mean weight-ratios judgments is an inverse one.

Table II summarizes the two analyses of variance performed on the log judgments of weight-ratio, one for Trial 2 and one for Trial 3. On

TABLE II

ANALYSIS OF JUDGMENTS: VARIANCE OF WEIGHT-RATIOS

(Log₁₀ plus 1.000 transformation)

Source	df.	Trial 2 *MS*	Trial 2 *F**	Trial 3 *MS*	Trial 3 *F**
Velocity-ratios (*VR*)	4	1.4565	7.10§	1.3327	5.08‡
Size-ratios (*SR*)	4	4.4682	22.32§	3.8987	19.59§
*S*s	18	.0999	.49	.1359	.59
VR×*SR*	16	.0349	2.49‡	.0308	2.75‡
VR×*S*s	72	.2052	14.66§	.2626	23.45§
SR×*S*s	72	.2002	14.30§	.1990	17.77§
Residual	288	.0140		.0112	

* For interactions, *MS* tested against residual *MS*. *MS*s of *VR* and *SR* tested respectively against *VR*×*S*s and *SR*×*S*s. *MS* of *S*s tested against pooling of *VR*×*S*s and *SR*×*S*s.

‡ Significant at 1% level.

§ Significant at ⅒% level.

both trials the velocity-ratios and the size-ratios are highly significant in their effects on the log judgments of weight-ratio. Size, however, seems to be the greater determinant. Plots (not shown here) of the geometric means relevant to the *VR* × *SR* interaction show no systematic difference from level to level.

Discussion

A disparity between certain present findings with respect to judgments of velocity and some earlier research requires comment. Fig. 1B shows that no relationship was found between the overall effect of the size-ratios employed and the geometric mean velocity-ratio judgments found.

Brown did find a relationship between physical size and velocity.[9] He worked with two equal-sized apertures so placed that only one could be seen at a time; viewing the other required turning the head. Behind one aperture and clearly visible

[9] J. F. Brown, The visual perception of velocity, *Psychol. Forsch.*, 14, 1931, 199-232.

through it passed an equally spaced array of circles, 0.8 cm. in diameter. In the other aperture could be seen another equally spaced array of circles of 1.6 cm. diameter. The *Ss* had to adjust the velocity of the latter array until it equaled the velocity of the standard series (0.8 cm.). The larger figures were judged as moving more slowly than the smaller ones. A diameter relation of 1:2, it was found, requires a mean velocity ratio of 1.25:1 before the *Ss* judge the two arrays as moving equally fast. Brown also had *Ss* compare arrays of circles whose diameters were three and four times the size of those of the standard series. The velocity-ratios were set on the average at 1.32 and 1.39 to one, respectively.

When the diameters of two sets of figures are in the ratio of 1:2 in Brown's study, the areas are in the ratio of 1:4. It may be that the ratio of sizes must be beyond a certain value before it can have an effect on judged velocity. Even with the ratio of sizes at 1:4, it was found that the velocity settings were influenced only to the extent of 1.25:1. Perhaps the present study did not involve sufficiently great differences in the areas of the two colliding figures to produce any significant effect. Another possible explanation for the disparity in results may be that Brown's method is of greater sensitivity. Setting velocities may pick up small effects which ratio-statements do not.

On the other hand, both the size of a figure and its velocity influence the judgment of weight. An inverse relationship seems to exist with respect to velocity and a direct one with respect to the size of the figure.

A by-product of the present study is that it makes available a large set of judgmental values of velocity and weight. A study of the kind Michotte has done can now be conducted and mean estimates of the velocity- and weight-ratios may be used as predictive variables.

Summary

A psychophysical experiment was performed in which two groups of 19 *Ss* responded on three trials to the same set of 25 stimulus-objects, in three different counterbalanced orders. In all trials a black square or rectangle (*A*) move to make contact with a red square or rectangle (*B*). When this occurs, *B*, previously stationary, moves off in the same straight line as *A* while *A* remains stationary. The collisions varied with respect to both the size- and velocity-ratios of *A* to *B*. For both dimensions, ratios of 3:1, 2:1, 1:1, 1:2, and 1:3 were used. The *Ss* of Group I estimated the velocity-ratio of each stimulus-condition; those of Group II estimated the size-ratio.

The following results were obtained from the *Ss*' judgments: (a) an inverse relationship between physical velocity-ratios and geometric mean weight-ratios, (b) a direct relationship between the physical size-ratios and the geometric mean weight-ratios; (c) a direct relationship between the physical velocity-ratios and the geometric mean velocity-ratios; and (d) no relationship between the physical size-ratios and the geometric mean velocity-ratios.

TEXTURE-GRADIENTS AND JUDGMENTS OF SLANT AND RECESSION

By JACOB BECK, University of Pennsylvania

Gibson, Purdy, and Lawrence investigated the qualities of slant and recession produced by peripheral-to-central texture-gradients present in the cross-sections of the ray-sheafs stimulating the two eyes.[1] They reported on the judgments of slant and recession produced by a zero gradient and by an increasing density-gradient whose rate of increase was equal to that present in the cross-section of rays when looking through an 'equally striped' black and white parallel-sided tunnel. This study suggested that a texture-gradient without supplementary motion or disparity determines only a certain *family* of surfaces in depth, the gradient being ambiguous with respect to members of this family. Gibson had formerly argued that a texture-gradient as such determines a certain impression of surface-recession and slant.[2] There was, however, evidence to suggest that the ambiguity was eliminated when the texture-gradient was supplemented by one of disparity or of motion-parallax.

The purpose of the present experiment was: (1) to extend the results of the first study to other texture-gradients; and (2) to determine whether binocular disparity does in fact eliminate ambiguity of apparent slant and recession. Judgments of slant and recession produced by four texture-gradients, whose cross-sections were equivalent to those of non-cylindrical objects, were studied. The gradients could be viewed either monocularly, with a motionless head, or binocularly.

METHOD. The method for controlling and systematically varying texture-gradients has been described in detail.[3] A pseudo-tunnel (an optical tunnel) is constructed by placing black and white plastic sheets, from whose centers circular holes have been cut, alternately behind one another on a pair of parallel steel tracks 21 ft. long. The centers of the holes are at eye-level for a seated *O*. The distances of the sheets from *O* determine the angular size of the black and white transitions in the cross-

* Received for publication July 31, 1958. This research was supported in part by the Office of Naval Research under Contract Nonr 401 (14) with Cornell University, under the direction of Professor James J. Gibson.

[1] J. J. Gibson, Jean Purdy, and Lois Lawrence, A method of controlling stimulation for the study of space perception: The optical tunnel, *J. exp. Psychol.*, 50, 1955, 1-14.

[2] Gibson, *The Perception of the Visual World*, 1951, 85-94.

[3] Gibson, Purdy, and Lawrence, *op. cit.*, 1-14.

section of rays. Thus, by changing the spacing of the sheets along the tracks, the peripheral-to-central gradients of texture-density in the cross-sections can be varied systematically. Uniform lighting was produced by fluorescent lamps situated beneath and running the length of the optical tunnel. The light was diffused by several thicknesses of tissue paper. Fig. 1 illustrates the arrangements employed here.

Experiment I. An 11-element tunnel was constructed. The distances between suc-

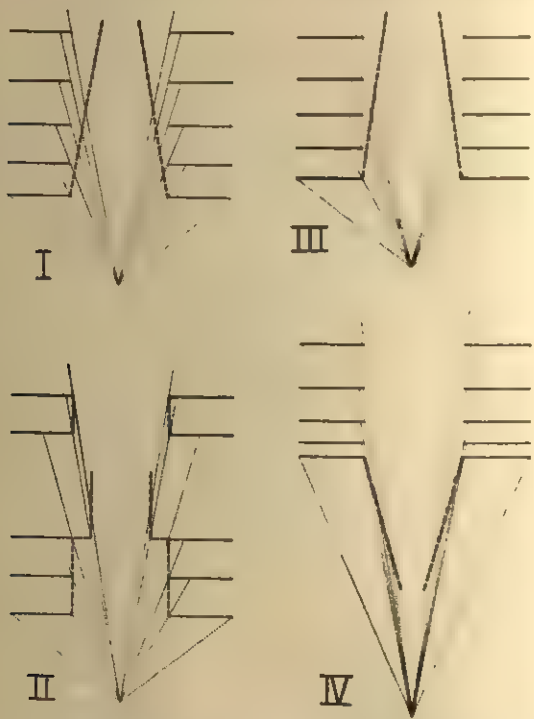


FIG. 1. CONDITIONS OF THE FOUR EXPERIMENTS

Five elements, out of a large number used, are shown to illustrate the spacing of the plastic sheets. The broken line indicates the reflecting surface with equally-spaced stripes whose margins coincide with the margins of the sheaf of rays serving as the optical stimulus. In all diagrams, the rays intercept equal segments along the broken line. Note that for Experiment II two such surfaces are shown: two parallel-sided tunnels of the same diameter separated by a gap, and a large parallel-sided tunnel followed immediately by a parallel tunnel of smaller diameter.

cessive plastic sheets were set to duplicate the increasing density-gradient present in the cross-section of rays when looking through the wide end of an equally-striped, tapered tunnel which converges to an angle of 20°.

Experiment II. A 10-element tunnel was constructed to produce a sudden change in density of the ray-sheaf, but with the rate of increase constant on either side of the change. The plastic sheets were placed at equal distances, except for a gap which occurred between the fifth and sixth sheets.

Experiment III. A 19-element tunnel was constructed which duplicated the increasing density-gradient present in the cross-section of rays when looking through the wide end of an equally-striped, tapered tunnel which converges to an angle of 16°.

Experiment IV. By greatly increasing the spacing between successive plastic sheets, an 11-element tunnel was constructed with a peripheral-to-central decrease of density duplicating the gradient which occurs when looking at the tapered end of an equally-striped, truncated cone which converges to an angle of 32°.

The stimulus in the experiments was the sheaf of light-rays determined by the spacing of the plastic sheets along the tracks. In Experiments I, III, and IV, the margins of only one reflecting surface with equally-spaced stripes coincided with the margins of the sheaf of light-rays serving as the optical stimulus. In Experiment II, the margins of many different reflecting surfaces with equally-spaced stripes coincided with the margins of the sheaf of light-rays. As Fig. 1 shows, two parallel-sided tunnels of the same diameter separated by a suitable gap, or a large tunnel followed immediately by a tunnel of smaller diameter, or any pair of tunnels of

TABLE I

TABULATION OF *O*'S REPORTS IN EXPERIMENT I

Mode of observation	Solid		Form		Stripes	
	yes	no	tapered	parallel	equal	unequal*
Monocular	13	2	14	1	15	0
Binocular	10	5	0	15	0	15

* Unequal stripes becoming wider as they recede.

intermediary diameter and separation would have equally-spaced stripes and would coincide with the margins of the sheaf of light-rays serving as the optical stimulus.

In Experiments I and II, the *O*s viewed the texture-gradients first with fixed head and monocular vision, and then binocularly. The effect of binocular vision was to add a positive disparity-gradient which indicated the true slant and recession of the tunnels. In Experiment III and IV, *O*s viewed the stimuli: (a) binocularly through a prism-pseudoscope; (b) monocularly but through one tube of the pseudoscope; and (c) monocularly without the pseudoscope. A prism-pseudoscope reverses the normal gradient of disparity and would thus be expected to reverse the qualities of slant and recession.

Observers and instructions. The *O*s were naïve with regard to the experimental situation and its purpose. The instructions were as follows:

> When I lift the panel in front of you, you will see a black-and-white striped object which I would like you to describe to me. Please look at it so carefully that afterwards you will be able to tell me as much about it as possible. Please note such things as the shape and dimensions of the object you see. You may compare it to something you may have seen before, if you like.

Specific questions about the spacing of the stripes, the solidity, and the shape of the object *O* saw were asked if the information had not been included in his free description. *O* was allowed to observe the optical tunnel for only 10 sec. at a time, but the number of observations was not limited. Generally, no more than three observations were necessary to answer the questions asked.

The *O*s in Experiments I and II were the same, as were the *O*s in Experiments

III and IV. A balanced order was, therefore, employed with half the *O*s viewing each optical tunnel first. In Experiments I and II, 15 *O*s were used, and in Experiments III and IV, 12 *O*s were used.

Results. The questions posed are these: Is a texture-gradient in isolation a sufficient stimulus for producing an impression of slant and recession? Is this impression variable or not? What are the effects of adding binocular disparity?

Tables I, II, and III reflect the *O*s' descriptions in each of the four experiments. Inspection of the tables indicates that a texture-gradient in isolation is able to evoke complex impressions of slant and recession. With fixed monocular vision, the texture-gradients in Experiments I, III, and IV produced the impressions of non-cylindrical tapered objects which the

TABLE II

TABULATION OF *O*S' REPORTS IN EXPERIMENT II*

Mode of observation	Solid		Gap		Diameter of tunnels	
	yes	no	yes	no	2nd smaller	both equal
Monocular	12	3	4	11	15	0
Binocular	12	3	15	0	6	9

* Both with monocular and with binocular vision, all the *O*s reported the two tunnels as being parallel sided and having stripes of equal width.

*O*s described by the terms "funnel," "narrowing pipe," or "narrowing tunnel." The texture-gradient in Experiment II produced the impression of a large cylinder followed by a second cylinder of smaller diameter which for the majority of the *O*s appeared directly connected to the first cylinder with no gap between.

The tables show that the surfaces and recessions perceived were variable. The impressions were most uniform in Experiment I, differed in the apparent gap separating the two cylindrical tunnels in Experiment II, and differed in the perceived convergence of the tunnels in Experiments III and IV. We must conclude that slant and recession are not uniquely determined under the reduced stimulus-conditions of fixed monocular vision.

The variable impressions of slant and recession occurring with a single ray-sheaf are not surprising when one remembers that a sheaf of rays may have for its reflecting surface any surface whose margins coincide with the margins of the ray-sheaf. An additional hypothesis is, therefore, required to account for the arousal of an impression of slant and recession by a fixed monocular stimulus. The readiest hypothesis would be that it is determined by a presumption, attitude, or expectation. One example of a presumption would be that the object is a cylinder, and another would be

that the stripes of the surface are equally spaced. In Experiments I and II, the *Os* showed a strong presumption to see an even scatter of texture, with all *Os* reporting the stripes as being equally spaced throughout the tunnels. In Experiments III and IV, this did not occur, and all but three *Os* reported the stripes of the tunnels as unequal in width. In these experiments, the failure for a presumption of an even scatter of texture to occur may have been due to the expectancy of an uneven texture established by previously having viewed the stimuli binocularly through the pseudoscope. This together with the delicacy of the stimulus-situation may also have been the reason for the failure of the texture-gradient to produce the impression of a convex cone with equally spaced stripes in Experiment IV.

To define the variability occurring with a single ray-sheaf, a linkage between judgments of slant and judgments of texture-spacing has been sug-

TABLE III

TABULATION OF *Os*' REPORTS IN EXPERIMENTS III AND IV*

Mode of observation	Experiment III				Experiment IV			
	flat	concave	convex	concave-convex	flat	concave	convex	concave-convex
Binocular with pseudoscope	1	4	6	1	0	8	3	1
Monocular with pseudoscope	1	6	4	1	0	10	2	0
Monocular	0	11	1	0	0	11	1	0

* In both experiments, all the *Os* reported the tunnels as appearing solid.

gested.[4] The hypothesis proposes that, with a presumption of texture-spacing, the apparent surface-slant must be that physical slant which would yield the current retinal image; with a presumption of surface-slant, the apparent texture-spacing must be that physical spacing which would yield the retinal image. Thus, a further question is: Does the covariation of texture-spacing and slant which together combine to yield a particular ray-sheaf determine a unique coupling of apparent slant to apparent texture-spacing?

A test of this hypothesis would have required independent quantitative judgments by *O* of tunnel-convergence and texture-spacing. A linkage between apparent slant and apparent texture-spacing is indicated, however, by the change of *O*'s judgments from an equally-striped, tapered tunnel to an unequally-striped, parallel tunnel with the addition of binocular parallax in Experiment I. Furthermore, in accordance with the hypothesis, the *Os* of Experiment III reported the stripes as becoming wider when receding if they reported the tunnel to be 'deep' and almost parallel and narrower when receding if they reported the tunnel to be 'shallow.' Judgments were obtained, however, which indicate that this linkage is only approximate. For ex-

[4] Gibson, Purdy, and Lawrence, *op. cit.*, 10-11.

ample, one O in Experiment I reported the tunnel to be both parallel-sided and possessing an equal spacing of texture; in Experiment III, an O viewing the tunnel binocularly through the pseudoscope reported it as appearing flat with the stripes becoming narrower toward the periphery—according to the hypothesis, they should appear to become wider toward the periphery; in Experiment IV, a number of similar exceptions occurred. We must conclude, therefore, that the family of perceptions of slant and recession which may be evoked by a texture-gradient in isolation corresponds only approximately with the set of surfaces defined by the margins of the ray-sheaf.

What was the effect of binocular disparity? Adding a positive disparity-gradient in Experiment I and II and a negative disparity-gradient in Experiment III and IV changed O's reports of slant and recession in the direction of the gradients. In Experiment I, the gradient of disparity won out over a conflicting gradient of texture-density (assuming a presumption for an equal spacing of stripes), and all the Os reported a parallel tunnel with the stripes becoming wider as they receded. In Experiment II, the gradients of disparity and texture were not discrepant, and binocular vision simply changed all reports to two cylinders behind one another with a gap between. McNemar-tests of the differences in the number of Os reporting parallels vs. tapered in Experiment I and gap vs. no gap in Experiment II were significant at or beyond the 1% level. In Experiments III and IV, a negative disparity-gradient lessened the number of concave judgments when added to conflicting or non-conflicting texture-gradients (1% level in I and 10% level in II). The hypothesis that the ambiguity of fixed monocular vision is eliminated when binocular parallax is operative was not, however, supported. The experiments showed again that the way separate stimuli interact is complex. In Experiments III and IV, viewing the stimuli through the pseudoscope actually increased the variability of the Os' reports.

Summary

By means of an optical tunnel, four different texture-gradients were studied. The results show that, with fixed monocular vision, a texture-gradient alone can evoke a complex impression of slant and recession. The qualities of slant and recession are not, however, uniquely determined. The hypothesis that the ambiguity of a texture-gradient is only with respect to the members of a geometric family of surfaces must be rejected, since the linkage between judgments of texture-spacing and surface-slant is only approximate. Adding a disparity-gradient to the texture-gradient tends to change the Os' judgments of slant and recession in the direction of the disparity-gradient. The disparity-gradient does not, however, eliminate all the ambiguity of fixed monocular vision.

FIXED-INTERVAL AND FIXED-RATIO PERFORMANCE IN THE FISH AS A FUNCTION OF PREFEEDING

By Rochelle M. Eskin and M. E. Bitterman, Bryn Mawr College

In a series of experiments under way in this laboratory, the behavior of the fish *Tilapia macrocephala* is being studied in a variety of situations analogous to those used for the study of learning in the rat. Our orientation in this work is comparative; the experiments are designed to determine whether a single theory of learning will fit both forms, or whether different theories will be necessary.[1] Since the manipulation of drive plays so important a role in research on the rat, it is inevitable in such a program that the problem of manipulating drive in the fish must soon arise, and it is that problem—specifically, the problem of controlling hunger in the fish—which is considered here.

In one recent experiment with *Tilapia,* time of deprivation was varied, with results quite like those obtained for the rat; an orderly, negatively accelerated relation appeared between the amount of food required to satiate and time since a previous satiation.[2] In another experiment, total daily intake was varied, again with results like those obtained for the rat; a group of fish on a ration of 10 pellets per day showed greater resistance to extinction than did a second group on a ration of 30 pellets per day.[3] In the experiments reported here, the feasibility of prefeeding as a method of controlling hunger in the fish is studied. Haralson already has reported an effect of partial satiation on resistance to extinction in continuously reinforced *Tilapia.*[4] Here we study the effect of amount of prefeeding on fixed-interval and fixed-ratio responding as well as on resistance to extinction after fixed-interval and fixed-ratio training.

---

* Received for publication August 17, 1959. This work was supported by Grant M-2857 from the National Institute of Mental Health. We are indebted to Dr. Lester R. Aronson, of the American Museum of Natural History, for the animals used.

[1] M. E. Bitterman, Jerome Wodinsky, and D. K. Candland, Some comparative psychology, this Journal, 71, 1958, 94-110.

[2] Bitterman, Wodinsky, and Candland, *op. cit.,* 108-110.

[3] Nicholas Longo and M. E. Bitterman, The effect of partial reinforcement with spaced practice on resistance to extinction in the fish, *J. comp. physiol. Psychol.,* in press.

[4] John Haralson, The effects of drive level on performance and extinction of a learned response in rats and fish, *J. comp. physiol. Psychol.,* 51, 1958, 732-736.

Method

Subjects. The *Ss* were eight female, sexually mature African mouthbreeders, each about 2.5 in. long, drawn from the stock tanks of the Department of Animal Behavior of the American Museum of Natural History.

Apparatus. A detailed description of the apparatus is provided elsewhere.[5] The manipulandum was a target at which *S* was trained to strike. The target was a fold of aluminum window-screening mounted on a light rod which was inserted into the needle-holder of a crystal phonograph-cartridge. The output of the cartridge was led to a high-gain amplifier, and the output of the amplifier energized a system of relays which determined the consequences of response. Response and reinforcement activated a Davis cumulative recorder. Reinforcement was a pellet of food (Aronson's mixture) dropped into the water by a mechanical feeder.[6] Each time the feeder operated, an overhead lamp dimmed momentarily.

Preliminary training. In the first stage of this work, the *Ss* were adapted to the experimental situation. Each *S* was carried to the apparatus in its individual living-tank. The target was not introduced, but at intervals of 1 min. a timing device caused the overhead lamp to dim and the feeder to drop a pellet of food into the water. Ten reinforcements per day were given until the food was being taken readily (which required about three days, on the average), and then training to the target was begun: the target, baited with liver, was lowered into the tank, and each contact of *S* with the target was reinforced with a pellet of food until 30 pellets had been earned. In subsequent sessions, the unbaited target was used, again with continuous reinforcement and 30 reinforcements per day. Training with the unbaited target was continued until a stable rate of response and a minimal initial latency were achieved (on the average, for another three or four days).

The *Ss* then were assigned at random to two different subgroups of four animals each, a fixed-interval (*FI*) group and a fixed-ratio (*FR*) group. Two of the four *FI Ss* were trained on a 2-min. schedule, one on a 1-min. schedule, and one on a 4-min. schedule. Two of the four *FR Ss* were on a 1:10 schedule, and two were on a 1:20 schedule. Each *S* was run every third day for whatever time was required to earn 20 reinforcements, and on non-training days each *S* was given 20 pellets in its home-tank at corresponding times. Following 7–11 such sessions, all *FI Ss* were trained for 5 sessions on a 2-min. schedule, and all *FR Ss* were trained for 5 sessions on a 1:20 schedule.[7] The animals were trained every other day on a maintenance schedule of 40 pellets each training day, of which 10 were earned in the experimental situation.

Experiment I. In this experiment, the effects of prefeeding 1, 15, or 30 pellets were compared, with each *S* serving as its own control. The *FI Ss* were continued on a 2-min. schedule and the *FR Ss* on a 1:20 schedule. Each *S* was tested every

---

[5] Longo and Bitterman, Improved apparatus for the study of learning in the fish, this Journal, 72, 1959, 616-620.

[6] The pellets used were of an attractive dried food, the recipe for which may be found in L. R. Aronson, An analysis of reproductive behavior in the mouthbreeding cichlid fish, *Tilapia macrocephala* (Bleeker), *Zoologica*, 34, 1949, 136 (footnote). The pellets were screened for uniformity of size, their largest dimension being 1.5-2.0 mm.

[7] One of the *FI* animals, previously trained on a 1-min. schedule, was given a single 12-hr. period of *FI* training on an 18.3-min. schedule before being shifted to the 2-min. schedule.

other day for as long as required to earn 10 pellets. At the beginning of each experimental session, *S* received either 1, 15, or 30 pellets, and, 1 hr. after the session, either 29, 15, or no pellets (for a total of 40); no food was given on nonexperimental days. There were seven training sessions and four balanced patterns of prefeeding, each *S* of each group being assigned at random to one of the four patterns. In each session, a tracing of cumulative response was made, and the time between the first response and the last reinforcement was measured with a stopwatch.

Experiment II. In this stage of the work, the *Ss* were given 31 pellets every day, and a week elapsed between the end of the first experiment and the beginning of the second during which they were given an opportunity to adjust to the new regimen. Each *S* was tested on alternate days, with the schedules of reinforcement as before. There were two levels of prefeeding, 1 or 20 pellets, with 11 pellets earned in the experimental session, and 19 or no pellets given one hour afterward. (In this, and in the following experiment, the first response to the target in each training session

TABLE I

PATTERNS OF PREFEEDING IN EXPERIMENT III

Pattern	Session 1 (Tr)	2 (Ex)	3 (Tr)	4 (Ex)	5 (Tr)	6 (Ex)	7 (Tr)	8 (Ex)
1	1	30	20	1	1	30	20	1
2	20	30	1	1	20	30	1	1
3	1	1	20	30	1	1	20	30
4	20	1	1	30	20	1	1	30

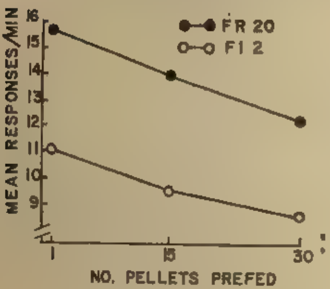
was reinforced.) There were four training sessions for each *S*, two *Ss* in each group being prefed in the sequence 1, 20, 20, 1, and two in the sequence 20, 1, 1, 20. In this experiment, the latency of the first response (the time between the insertion of the target and the occurrence of the first response) was measured, as well as the time between the first response and the last reinforcement.

Experiment III. As in Experiment II, each *S* received 31 pellets per day and was tested every other day; but now training sessions (in which 11 reinforcements were earned) alternated with extinction sessions (30 min. of unreinforced responding). In all, there were eight sessions, four of each kind. In the training sessions, either 1 or 20 pellets were prefed, and, in the extinction sessions, either 1 or 30 pellets were prefed, as shown in Table I. There were four different patterns of prefeeding, one *S* in each group being randomly assigned to each. Again in this experiment, the latency of the first response and the time from the first response to the last reinforcement were measured.

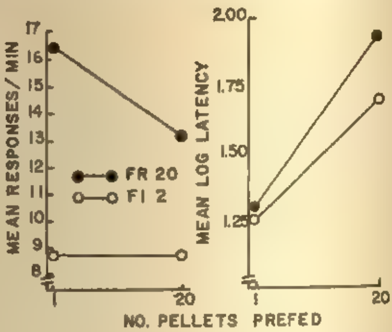
RESULTS

The results of the three experiments taken together show a clear inverse relation between amount prefed and rate of reinforced *FR* responding. The results for Experiment I, shown in Fig. 1, are significant beyond the 5% level of confidence by Friedman's test. Because of the small number of *Ss* and conditions used, no meaningful statistical tests can be made of the results of Experiments II and III, but the data are highly consistent from *S* to *S*

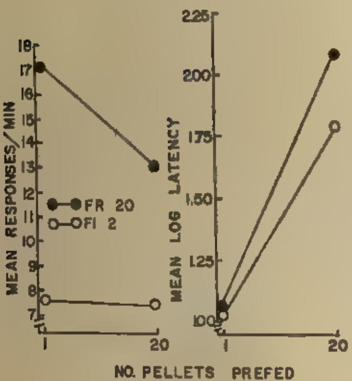
and from experiment to experiment (see Figs. 2 and 3). Rate of reinforced *FI* responding does not seem to be influenced by amount of prefeeding; the effect suggested by Fig. 1 does not achieve significance at the



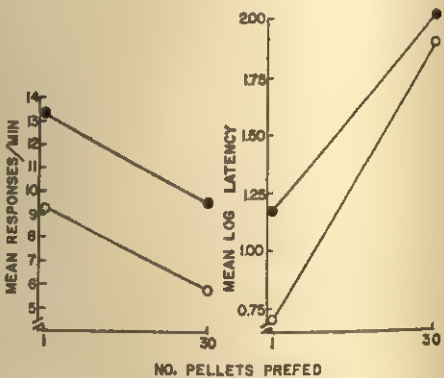
1



2



3



4

Fig. 1. Rate of Response as a Function of Prefeeding in Experiment I

Fig. 2. Rate of Response and Latency of Response as a Function of Prefeeding in Experiment II

Fig. 3. Rate of Response and Latency of Response as a Function of Prefeeding During the Reconditioning Sessions of Experiment III

Fig. 4. Rate of Response and Latency of Response as a Function of Prefeeding During the Extinction Sessions of Experiment III

5% level, and no effect at all is suggested by Fig. 2 or Fig. 3. The extinction data of Experiment III (Fig. 4) do, however, show parallel relations between rate and amount prefed for the two groups. Here the data for the *FI* group are every bit as consistent as for the *FR* group, there being not a single reversal in either case. In general, these results resemble those which have been obtained in work with higher animals—pigeon and rat

—which suggest that rate of response in Skinnerian situations increases with drive as manipulated by a variety of methods.[8] Apparently, substantial control of drive level in the fish may be achieved by prefeeding. The present results differ from those for higher animals only in that performance in *FI* training seems to be less sensitive to changes in drive.

Two qualitative findings may be worth noting. The first is that the records of the *FI Ss*, except for a lower over-all rate of response, are not readily distinguishable from those of the *FR Ss*; brief pauses after reinforcement are common, but no indication of scalloping appears. It cannot be concluded, however, that this absence of scalloping in the fish reflects some functional difference between fish and higher animals. Whether scalloping can be depended upon to appear spontaneously in higher animals —which once was taken for granted—seems now to be in doubt.[9] The second qualitative finding concerns the basis for the over-all change in the rate of *FR* responding produced by prefeeding. Sidman and Stebbins assert that prefeeding does not affect the rate of *FR* responding *per se*, but only the number and duration of periods of non-responding, which occur primarily after reinforcement.[10] This generalization, based on work with rat, cat, and monkey, is confirmed for the pigeon by Ferster and Skinner.[11] The records for the fish fail to support this generalization; the change in over-all rate produced by prefeeding can by no means be accounted for entirely in terms of an increase in the number and duration of periods of non-responding. It is possible, of course, that a more stable rate of response, less susceptible to the effects of prefeeding, might have been established with prolonged *FR* training.

Qualitative observations made in Experiment I suggested that latency of the first response to the target also might be sensitive to the effects of prefeeding. This impression was confirmed by the results of Experiments II and III (Figs. 2, 3, and 4), which show a consistent positive relation

---

[8] B. F. Skinner, Conditioning and extinction and their relation to drive, *J. gen. Psychol.*, 14, 1936, 296-317; B. R. Horenstein, Performance of conditioned responses as a function of strength of hunger drive, *J. comp. physiol. Psychol.*, 44, 1951, 210-224; N. E. Miller, Shortcomings of food consumption as a measure of hunger; results from other behavioral techniques, *Ann. N.Y. Acad. Sci.*, 63, 1955, 141-143; Bernard Weiss and Edwin Moore, Drive level as a factor in distribution of responses in fixed-interval reinforcement, *J. exp. Psychol.*, 52, 1956, 82-84; C. B. Ferster and B. F. Skinner, *Schedules of Reinforcement*, 1957, 71-72, 320; F. C. Clark, The effect of deprivation and frequency of reinforcement on variable-interval responding, *J. exp. anal. Behav.*, 1, 1958, 221-228.

[9] Skinner, Some contributions of an experimental analysis of behavior to psychology as a whole, *Amer. Psychologist*, 8, 1953, 69-78; Weiss and Moore, *op. cit.*, 82-84.

[10] Murray Sidman and William Stebbins, Satiation effects under fixed-ratio schedules of reinforcement, *J. comp. physiol. Psychol.*, 47, 1954, 114-116.

[11] Ferster and Skinner, *op. cit.*, 71.

between mean log latency and amount prefed, under *FI* as well as under *FR* conditions. (In Haralson's experiment, too, latency was increased by prefeeding in the fish, although a comparable effect did not appear in the rat.)[12] The curves for the *FI* group suggest that latency is more sensitive than is rate to changes in drive. In any event, the results for latency, like those for rate, demonstrate that control of drive in the fish may be achieved by prefeeding.

The results of Experiment III bear also on the adjustment of *Tilapia* to repeated conditioning and extinction. In work with the rat, level of per-

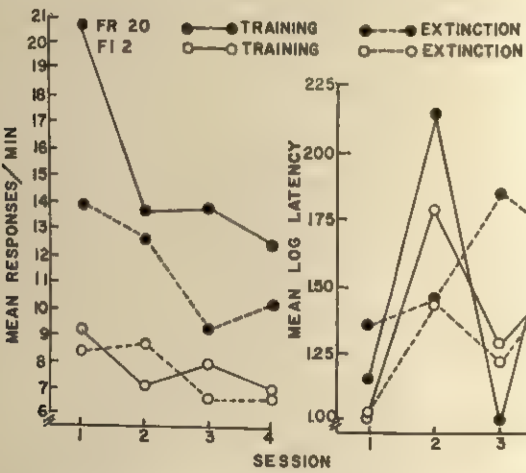


FIG. 5. CHANGES IN RATE OF RESPONSE AND IN LATENCY OF RESPONSE FROM SESSION TO SESSION IN EXPERIMENT III

formance in successive reconditioning sessions has been found to increase while level of performance in interpolated extinction sessions progressively declines,[13] and similar results have been found for the fish in an experiment with discrete trials.[14] In the present experiment, with rate as the measure, performance during reconditioning and resistance to extinction both tend to decline as a function of sessions (Fig. 5). In fact, the reconditioning decrement is the more reliable of the two; that for the *FR* group is significant beyond the 2% level by Friedman's test, while none of the other decrements approaches statistical significance. The latency data show no sig-

[12] Haralson, *op. cit.*, 732-736.

[13] D. H. Bullock and W. C. Smith, An effect of repeated conditioning-extinction upon operant strength, *J. exp. Psychol.*, 46, 1953, 349-352.

[14] Wodinsky and Bitterman, Partial reinforcement in the fish, this JOURNAL, 72, 1959, 184-199.

nificant changes over sessions. Further studies of the adjustment of the fish to conditioning-extinction series perhaps will throw some light on these findings.

Summary

In three experiments with the fish *Tilapia macrocephala*, prefeeding reduced the rate of fixed-ratio responding but had no consistent effect on the rate of fixed-interval responding. The latency of the first response in each session increased with prefeeding both in *FI* and in *FR* training, and resistance to extinction decreased with prefeeding after training under both conditions. These results, analogous to those obtained with higher animals, suggest that substantial control of drive in the fish may be achieved by prefeeding.

SUBJECTIVE INTENSITY OF COFFEE ODOR

By T. S. Reese and S. S. Stevens, Harvard University

The aim of this study was to develop a method for presenting controlled intensities of olfactory stimuli derived from natural coffee, and to determine, by the method of magnitude-estimation, how the subjective intensity of this odor grows with stimulus-concentration.[1] Since the 'psychophysical law' relating psychological magnitude to physical magnitude on a wide variety of prothetic sensory continua has repeatedly turned out to approximate a power function, it was to be expected that the apparent strength of the coffee odor would be proportional to the concentration raised to a power.[2] Some earlier experiments with benzaldehyde, carried out with an apparatus not too well suited to the purpose, seemed to confirm the power function.[3]

Extensive research by Jones has also confirmed the power function for olfactory stimuli.[4] For nine organic compounds he found the value of the exponent of the power function to range from about 0.4 to about 0.6. The procedure used was simple sniffing from wide-mouthed bottles which contained various proportions of the organic substance mixed in oil.

Apparatus. To avoid some of the uncertainties that characterize efforts to control the exact concentration of an odor when it is sniffed from an open bottle, or is forced into the nose by 'blast injection,'[5] a technique was developed involving the use of 'sniffing-bags.'

Controlled concentrations of the odorous gas were placed in nonpermeable, collapsible containers, each fitted with a nose-piece through which *O* could inhale the

* Received for publication October 21, 1958. Research supported by the National Science Foundation and the Office of Naval Research [Contract Nonr-1866(15); Project Nr142-201; Report PNR-219].

[1] The method of magnitude-estimation is described in S. S. Stevens, The direct estimation of sensory magnitudes—loudness, this Journal 69, 1956, 1-25. For its relation to other psychophysical methods, see Stevens, Problems and methods of psychophysics, *Psychol. Bull.*, 54, 1958, 177-196.

[2] Stevens, On the psychophysical law, *Psychol. Rev.*, 64, 1957, 153-181.

[3] These results are reported in T. S. Reese, Magnitude estimation of smell, Honor's thesis in Psychology, Harvard University, 1957. See also Stevens, *op. cit.*, 1957, 166.

[4] F. N. Jones, Scales of subjective intensity for odors of diverse chemical nature, this Journal, 71, 1958, 305-310; Subjective scales of intensity for three odors, this Journal, 71, 1958, 423-425.

[5] Carl Pfaffmann, Taste and smell, in S. S. Stevens (ed.), *Hanbbook of Experimental Psychology*, 1951, 1162-1165; Jones, A comparison of the methods of olfactory stimulation; blasting vs. sniffing, this Journal, 68, 1955, 486-488.

contents of the bag in the course of a normal sniffing process. This technique has the advantages that only the gas in the bag enters the nostrils and the process of sniffing is effectively carried out under normal atmospheric pressure.

The problem of finding a suitable material for the bags was given considerable attention. The ideal material needs to be impermeable, free of odor, thin, and flexible. These requirements seem to be met reasonably well by a material made of plastic and aluminum ("Scotchpak" 20A20, manufactured by Minnesota Mining and Manufacturing Co.). If the Scotchpak is thoroughly cleaned with soap and water it can be rendered essentially odorless. It can readily be made into bags by heat-sealing the edges. For the present experiment, bags with a volume of 4 liters were made by heat-sealing the edges of two triangular pieces 15 in. across the base and 17 in. high. The neck of the bag was wrapped tightly around a 4-in. section of glass tubing of 7 mm. interior diameter.

A nose-piece was attached to the tubing in the neck of each bag. It consisted of two glass plugs, each about the size of the end of a thumb, and so positioned that the plugs could be seated tightly and comfortably over the entrances to the nostrils. The plugs were made of ground Pyrex ball-joints 18 mm. in diameter with a 7 mm. bore, and were fastened to a fork leading to the tube that led in turn to the bag. To make the ball-joints fit the typical nostrils, their axes were set at an angle of 60° toward one another. It was also necessary to grind flat the two facing surfaces of the balls to make room between them for the septum of the nose. The angle between these flattened surfaces was about 30°. The distance between the balls was about 5 mm.

The nose-piece was connected to the bag by a short section of polyethylene tubing. This tubing was closed off by a clamp which could be easily opened by *O*. The entire assembly of bag, nose-piece, and clamp could be handed to *O*, and he in turn could place it quickly to his nose, open the clamp and inhale the contents of the bag.

For the experiment with coffee, different concentrations of the odor were manufactured by mixing odor and air in various proportions. An appropriate amount of air 'saturated' with coffee was first put in the bag and then the bag was filled with pure air. The odorous part of the mixture was obtained by running air at a slow, uniform rate through a half-pound of ground coffee contained in a gas-washing bottle. The output of this bottle was filtered through a 1-in. pack of glass wool which served to trap the coffee dust. The amount of odor placed in a bag was metered by timing its flow into the bag. The rate of flow was held constant by means of a constant driving pressure, and the relative concentrations in the different bags was determined by the relative duration of the injection of the odor. Only the relative concentrations need be known to test whether subjective intensity grows as a power function of concentration and to determine the size of the exponent.

One disadvantage of the method was that, although the observations could be made in a matter of 10 min., the washing of the bags and the preparation of a new series of stimuli required from 1 to 2 hr.

Procedure. The method of magnitude-estimation was used to obtain a ratio-scale of the subjective intensity of the odor of coffee. Five concentrations covering a range of almost 100 to 1 were prepared, each concentration in a separate bag. The *O* was first told how to hold the bag to his nose, and how to open the valve and take a generous sniff of the contents. He then practiced by actually sniffing from a bag

containing a medium concentration of the odor. He was then told that he would be given the bags one by one and that his task was to estimate the intensity of the odor in each bag. The intensity in the first bag was to be called '10,' and *O*'s task was to assign numbers to the various intensities proportional to their apparent strength.

The stimulus-odor called '10' was the middle concentration of the series. Since the bags were kept behind a screen, *O* did not know how many different intensities were to be judged. Actually, each of the five concentrations was judged twice by each *O*. Between the first and second runs, the standard was presented a second time. The order of presentation of the stimuli was different for each *O*. An experi-

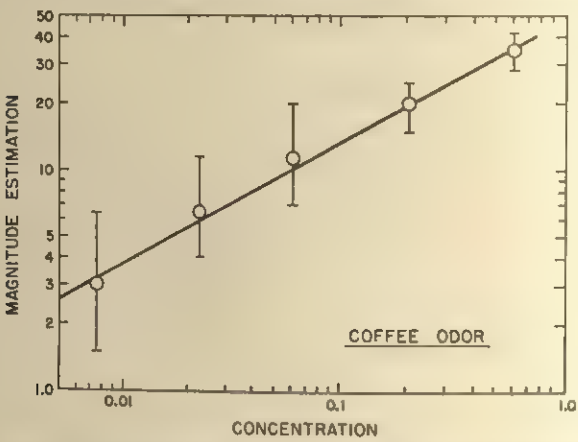


Fig. 1. Median Magnitude Estimations of the Subjective Intensity of Coffee Odor as a Function of Concentration
(Vertical bars show the interquartile range.)

mental session lasted from 5 to 10 min., and 20 to 30 sec. elapsed between the presentation of successive bags.

Results and discussion. The medians of the magnitude-estimations by 12 men are shown in Fig. 1. It is clear that these results approximate a straight line in a log-log plot, and, therefore, that they confirm the power law. In this case the slope of the line, and hence the exponent, is about 0.55.

In addition to the 12 men, a group of 8 women made judgments. Their results also approximated a power function but the slope was of the order of 0.3. This lower slope was due mainly to the fact that the judgments of three of the women were almost random. One of the women said there seemed to be little or no difference among the stimuli. Although there may be a sex difference here, it is rather more plausible to assume that

this small sample happened to include a few women who are inconsistent estimators of apparent magnitude. As pointed out elsewhere, one occasionally encounters people who use numbers and make estimates in rather strange ways.[6]

The combined results of all 20 *O*s, both men and women, produce a power function with an exponent of about 0.45. Although this value is based on a larger set of data, it seems probable, on the present evidence, that the exponent determined for the men alone (Fig. 1) is more representative of what additional measurements will show. This, however, remains to be seen.

The vertical lines in Fig. 1 mark the interquartile ranges of the magnitude-estimations. The distribution of the judgments is such that the medians coincide fairly closely with the geometric means. This type of distribution appears to characterize many of the experiments that use the method of magnitude estimation.[7]

In the course of the work that led up to the experiments on coffee, several exploratory experiments were conducted. Two of these studies bear on the subjective scale for heptane, which Jones found to be a power function with an exponent of about 0.59. Using sniffing bags made of polyethylene, we ran an experiment on 18 *O*s and obtained an exponent of about 0.6. Since there were certain difficulties with this experiment, including a tendency for the polyethylene to become contaminated, the experiment was repeated on 10 *O*s with the Scotchpak bags. The exponent then turned out to be about 0.68.

In the present state of the difficult art of controlling the olfactory stimulus, these various experiments seem to confirm one another about as well as can be expected. They suggest that the subjective intensity of the odor of heptane grows as a power function of the concentration, and that the exponent is of the order of 0.6. On the basis of the present evidence, it appears that there is also a power function for coffee and that its exponent is probably a little smaller than the exponent for heptane.

Although a direct tie-up between these subjective scales and recordable physiological processes is not yet possible, it is interesting to note that nerve potentials recorded from the olfactory bulb of a rabbit seem to grow with concentration roughly in the same manner as apparent intensity. Mozell found that the integrated spike discharge "increased approximately

---

[6] Stevens, this JOURNAL, *op. cit.*, 1956, 1-25.

[7] See, for example, J. C. Stevens, Stimulus spacing and the judgment of loudness, *J. exp. Psychol.*, 56, 1958, 246-250.

as a negatively accelerated function of concentration."[8] Such negative acceleration is consistent with the fact that the exponent of the power function is less than one. Whether this correspondence is more than a coincidence is difficult to say.

The sensory transducer involved in olfaction seems to behave as a 'compressor' in the sense that its output is a decelerating function of stimulus input. In this respect it resembles the auditory and visual transducers.[9] It contrasts with certain other modalitites (*e.g.* electric shock and kinesthetic force) in which sensory magnitude grows as an accelerating function of stimulus magnitude. For these modalities the exponent of the power function is greater than 1.0. 'Compression' is an obvious advantage in vision and hearing where the sense organ must handle stimulus inputs covering ranges of billions to one, but this principle hardly seems applicable to olfaction whêre the effective stimulus-ranges commonly encountered are mostly of the order of hundreds to one.[10] Despite the lack of any obvious biological utility, the sensory epithelium seems to behave as a compressive system.

Olfaction forms an exception to the rough relation noted elsewhere, that sensory continua for which the effective stimulus ranges are shorter have larger exponents.[11] A curious and possibly significant fact about smell is its limited dynamic range—subjectively speaking, the strongest odors we usually encounter are not many times more intense than the faintest odors.

Summary

A 'sniffing bag' was developed to facilitate the presentation of controlled concentrations of odorous gas, and this technique was applied to the scaling of the subjective intensity of natural coffee odor by the method of magnitude estimation. For a group of 12 men the subjective intensity of coffee was found to grow as the 0.55 power of the concentration. A group of 8 women gave a function with a flatter slope, but it is suggested that this result is probably not representative.

Two exploratory experiments with heptane gave evidence of a power function with an exponent of the order of 0.6. The fact that the apparent intensity of odors grows as a power function of the concentration is consistent with the law that appears to govern other sense modalities.

[8] M. M. Mozell, Electrophysiology of olfactory bulb, *J. Neurophysiol.*, 21, 1958, 183-196.
[9] Stevens, Measurement and man, *Science*, 127, 1958, 383-389.
[10] Cf. Mozell, *op. cit.*, 183-196.
[11] Stevens, Cross-modality validation of subjective scales for loudness, vibration, and electric shock, *J. exp. Psychol.*, 57, 1959, 201-209.

RESISTANCE TO EXTINCTION IN THE FISH AFTER EXTENSIVE TRAINING WITH PARTIAL REINFORCEMENT

By JEROME WODINSKY, New School for Social Research, and
M. E. BITTERMAN, Bryn Mawr College

The effect of partial as compared with consistent reinforcement on resistance to extinction in the fish *Tilapia macrocephala* was studied in two recent experiments, the first under conditions of massed practice,[1] and the second under conditions of widely spaced practice.[2] The response of the animal in each case was to strike at a target introduced into its living tank; the measure of performance was latency of response; and the index of resistance to extinction was number of trials to the criterion of five successive failures to respond in 30 sec. In both experiments, resistance to extinction was *less* after partial than after consistent reinforcement, at least initially. When the animals of the first experiment were subjected to a series of extinctions separated by reconditioning sessions, the paradoxical effect found in higher forms gradually made its appearance; resistance to extinction declined progressively in both groups, more rapidly in the Consistent than in the Partial, until the resistance of the Consistent Group was less than that of the Partial Group. The present experiment was designed to inquire into the cause of this shift in relative resistance to extinction.

One possibility to be considered is that the process which makes for greater resistance to extinction in the partially reinforced rat also operates in our fish, but more slowly, its effects being felt only after a much larger number of training trials. A second possibility is that the process in question does not operate in the fish, and that the shift in relative resistance which appeared in our first experiment is attributable to the more rapid development in the Consistent Group of a discrimination between the events of training and extinction. From the first point of view, a larger

* Received for publication July 10, 1959. This research was supported by Grant No. M-2857 from the National Institute of Mental Health. We are indebted to Dr. Lester Aronson of the American Museum of Natural History for the animals studied.

[1] Jerome Wodinsky and M. E. Bitterman, Partial reinforcement in the fish, this JOURNAL, 72, 1959, 184-199.

[2] Nicholas Longo and M. E. Bitterman, The effect of partial reinforcement with spaced practice on resistance to extinction in the fish, *J. compar. physiol. Psychol.*, 53, 1960, 169-172.

number of training trials prior to the initial extinction might be expected to produce results like those obtained in work with the rat. From the second point of view, overtraining should not be expected to influence the direction of the initial difference in resistance to extinction. In the experiment to be reported, the Partial and Consistent Groups were given four times as much training before the initial extinction as were the groups of the first experiment—more training than the earlier groups were given over the entire conditioning-extinction series.

Method

Subjects. The *S*s were 24 sexually mature and experimentally naïve African mouthbreeders (*Tilapia macrocephala*), about 2 in. long, drawn from the stock tanks of the American Museum of Natural History.

Apparatus. A detailed description of the apparatus is provided in an earlier paper.[3] The target was a fold of aluminum window-screening mounted on a light rod which was inserted into the needle-holder of a crystal phonograph-cartridge. The output of the cartridge was led to a high-gain amplifier, and the output of the amplifier energized a system of relays which determined the consequences of response. When the target was introducd into *S*'s 2-gal. individual living tank (in which *S* was brought to the experimental situation), a Standard Electric timer was started. The response of the animal (nosing at or striking the target) stopped the timer, momentarily dimmed an overhead lamp, and activated a mechanical feeder which (on reinforced trials) dropped a pellet of food into the water near the target. (The sequence of events on unreinforced trials was exactly the same as on reinforced trials, except that—although the feeder operated—no food was dropped.) The introduction of the lever and its withdrawal were accomplished manually by *E*.

Procedure. In the first stage of the experiment, the animals were adapted to the experimental situation. Each *S* in its individual living tank was taken from the storage rack and carried to the sink, where accumulated debris was siphoned off and the original level of water restored. Then the tank was taken to the experimental locus. No target was used, but the overhead lamp dimmed periodically and the feeder discharged a pellet of food into the water. When these pellets were being taken readily, the second stage of the experiment began. The target, baited with liver, was introduced into the tank, and each contact of *S* with the target was reinforced with a pellet of food.

In the third stage of the experiment, an unbaited target was used. Each *S* was given 20 consistently reinforced, discrete trials per day, for a number of days which varied somewhat from *S* to *S* because of individual differences in rate of adaptation. In this stage of training, and in subsequent stages, the target was withdrawn after each response and reinserted a few seconds later to begin the next trial. At the conclusion of this stage of the work, the 24 *S*s were paired to form two matched groups of 12 *S*s each. The basis of pairing was two-fold: the number of days of Stage-3 training (which ranged from 6-16) was considered, as was the mean latency of response on the last five days of that stage.

---

[3] Longo and Bitterman, Improved apparatus for the study of learning in fish, this Journal, 72, 1959, 616-620.

The fourth stage of the experiment lasted 40 days. Each animal of each group was given 20 trials per day. The *Ss* of one group, designated at random as the Consistent Group, were reinforced on each trial, while the remaining *Ss*, comprising the Partial Group, were reinforced on only half of each day's trials. The sequence of reinforced (R) and unreinforced (U) trials on odd days followed 20 different Gellermann-orders,[4] and the complementary orders were used on even days. Thus, if the order for the first 10 trials of an odd day was RRRUURURUU, the order for the first 10 trials of the next day was UUURRURURR. For purposes of equating drive-levels, each *S* of the Partial Group was given 10 additional pellets outside the experimental situation during this stage of the training. The feeding procedure was as follows: Each *S* was on a schedule of 22 pellets per day. If 20 pellets were earned in the experimental situation, 2 additional pellets were given an hour later in the living quarters. If 10 pellets were earned, 12 additional pellets were given in the supplementary feeding.

In the fifth stage of the experiment, which lasted only a day, the animals were extinguished. (All the animals could not, of course, be extinguished on the same calendar day. The training was staggered in such a way that no more than 6 *Ss* were scheduled for extinction on any given day.) The procedure on each extinction-trial was exactly the same as on the unreinforced trials for the Partial Group in Stage 4, and each *S* was given as many trials as required to bring it to the criterion of five successive failures to respond in 30 sec. or less. If, on any trial, an animal failed to respond within 30 sec., the lever was withdrawn and the trial was terminated, a latency of 30 sec. being recorded. After five successive trials of this kind, work with the animal was terminated.

Results

A logarithmic transformation was used for the latency-data, and, because latencies of less than 1 sec. could be measured reliably, 1 sec. was added to each measurement for the purpose of avoiding negative values.

During acquisition, the latencies of the two groups reached asymptotes which were significantly different at the 1% level by Wilcoxon's nonparametric test for paired replicates.[5] The asymptotic score of the Partial Group was somewhat above 0.08, while that of the Consistent Group was somewhat less than 0.06. In the earlier experiment, after which the present one was patterned, the response of the Consistent Group at asymptote also was more rapid than that of the Partial Group.

In both experiments, too, the animals of the Partial Group developed a tendency to respond less rapidly on trials following reinforcement than on trials following nonreinforcement. The results for the present experiment are summarized in Fig. 1, which may be compared with Fig. 10 of the earlier paper.[6] For each group, on each day of training in Stage 4, the

[4] L. W. Gellermann, Chance orders of alternating stimuli in visual discrimination experiments, *J. genet. Psychol.*, 42, 1933, 206-208.

[5] Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures*, American Cyanamid Company, 1949, 1-16.

[6] Wodinsky and Bitterman, *op. cit.*, 196.

mean latency-score for trials following those on which the Partial Group was reinforced was subtracted from the mean latency-score for trials following those on which the Partial Group was unreinforced. (Accordingly, a negative difference-score means faster response after nonreinforcement than after reinforcement.) As Fig. 1 shows, the difference-scores for the Partial Group became progressively more negative over the 40-day training period, while the control scores for the Consistent Group continued to hover about zero. The question was raised in our report of the earlier experiment whether a like result there obtained might be explained by the

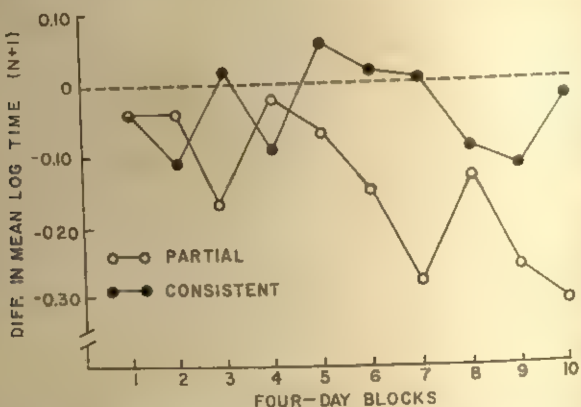


FIG. 1. DIFFERENTIAL RESPONSE IN THE PARTIAL GROUP AFTER REINFORCEMENT AS COMPARED WITH NONREINFORCEMENT AND CONTROL-DATA FOR THE CONSISTENT GROUP

The difference between mean latency-scores for trials following those on which the Partial Group was reinforced and mean latency-scores for trials following those on which the Partial Group was unreinforced is plotted in four-day blocks for the 40 days of training. A negative-score means faster response after nonreinforcement than after reinforcement.

fact that the probability of a reinforced trial after a nonreinforced trial was somewhat higher during training than the probability of a reinforced trial after a reinforced trial. This unlikely explanation is ruled out entirely by the results of the present experiment in which the Gellermann-orders were such as to exactly balance the two probabilities. As was pointed out in the earlier paper, partially reinforced rats trained in the runway show an analogous speed-differential when reinforced and unreinforced trials are regularly alternated, but not when Gellermann-sequences are employed.[7]

[7] D. W. Tyler, E. C. Wortz, and M. E. Bitterman, The effect of random and alternating partial reinforcement on resistance to extinction in the rat, this JOURNAL, 66, 1953, 57-65; E. J. Capaldi, The effect of different amounts of training on the resistance to extinction of different patterns of partially reinforced responses, *J. comp. physiol. Psychol.*, 51, 1958, 367-371.

Our principal interest in the outcome of the present experiment has to do, of course, with resistance to extinction. The performance of the two groups on the last day of training and on the first 70 trials of extinction is shown in Fig. 2, which may be compared with Fig. 1 of the earlier experiment.[8] The difference between the two groups on the last day of training is smaller in this experiment than in the first. During extinction, the curves converge as before in the early trials and then begin to diverge, the Partial Group tending to extinguish more rapidly than the Consistent Group. In the present experiment, this divergence begins sooner than it did before, being evident here after only 40 or 45 trials. In terms of trials to the criterion of five successive failures to respond in 30 sec. or less, the Consistent Group of the present experiment showed greater resistance to extinction than did the Partial Group. The mean of the Consistent Group was 230 trials, while that of the Partial Group was only 101 trials. The value for the Consistent Group as given is, in fact, spuriously low, because it includes the data of two animals whose behavior in extinction was clearly disrupted by power-failures that blacked out the laboratory. If these animals are excluded from

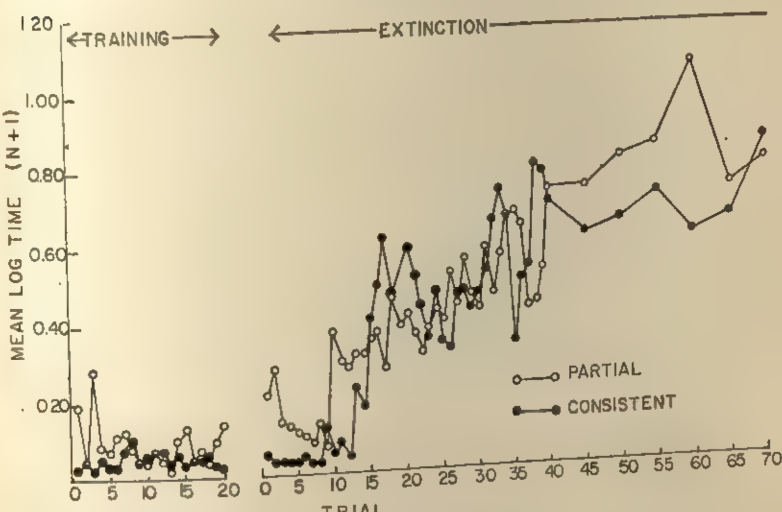


FIG. 2. PERFORMANCE ON THE FORTIETH DAY OF TRAINING AND ON TRIALS 1-70 OF EXTINCTION

After the fortieth trial of extinction, the latency-scores only for every fifth trial are plotted. Extinction continued well beyond the seventieth trial.

[8] Wodinsky and Bitterman, *op. cit.*, 189.

the computation, the mean of the Consistent Group becomes 251 trials. For 12 animals in each group, Wilcoxon's test for paired replicates makes it possible to reject the null hypothesis at the 6% level of confidence (two tails). The same test for 10 pairs (excluding the two disrupted fish of the Consistent Group and the Partial *Ss* matched with them) yields a confidence-level of 2% (two tails).

These results suggest that our earlier findings on initial resistance to extinction in partially and consistently reinforced fish are not simply to be attributed to insufficient initial training. The difference between the two groups is in the same direction and more pronounced, if anything, in the present experiment as compared with the first. Our findings support the conclusion, therefore, that the process responsible for the paradoxical effect of partial reinforcement which appears in mammalian experiments does not operate in our fish. The only indication of such an effect which has appeared thus far in our work with the fish is the shift in relative resistance to extinction of the partially and consistently reinforced *Ss* of our first experiment in the course of repeated conditioning and extinction, but that shift can be understood in terms of a difference in discriminability of the conditions of training and extinction. This interpretation, it may be well to note, contradicts the view of many investigators (however much their detailed treatments may differ) that the paradox of partial reinforcement and the phenomenon of adjustment to conditioning-extinction series are expressions of the same underlying process.[9] To be sure, we may continue to talk of the two effects in discriminative terms, but the levels of discrimination involved seem to be quite different. It is conceivable, for example, that the paradox of partial reinforcement reflects a capacity of the animal to respond to the properties of temporal series of events and so to achieve freedom from sheer frequency of occurrence.[10]

Summary

Two matched groups of African mouthbreeders were trained to strike at a target for food-reward. Then the animals were each given 800 discrete, massed trials (20 trials per day for 40 days) in which one group was consistently reinforced and the other partially reinforced. Despite prolonged training, the Partial Group showed less resistance to extinction than did the Consistent Group. The theoretical significance of this fact is noted.

---

[9] O. H. Mowrer and Helen Jones, Habit-strength as a function of the pattern of reinforcement, *J. exp. Psychol.*, 35, 1945, 293-311; C. C. Perkins, Jr. and A. J. Cacioppo, The effect of intermittent reinforcement on the change in extinction rate following successive reconditionings, *ibid.*, 40, 1950, 794-801; D. W. Lauer and W. K. Estes, Successive acquisitions and extinctions of a jumping habit in relation to schedule of reinforcement, *J. comp. physiol. Psychol.*, 48, 1955, 8-13.

[10] Tyler, Wortz, and Bitterman, *op. cit.*, 57-65; Capaldi, *op. cit.*, 367-371.

STANDARD STIMULUS-CONDITIONS FOR THRESHOLDS OF APPARENT MOVEMENT

By RAYFORD T. SAUCER and NEIL W. COPPINGER, VA Center, Kecoughtan, Virginia

As recent experiments have shown, the perception of apparent movement is affected by damage to the central nervous system,[1] by schizophrenia,[2] by chemotherapy,[3] and by many other conditions that have no known relationship to the central nervous system. It may well be, therefore, that the sheer process of aging may affect this perception and that useful information concerning aging and functional efficiency may be obtained through tests of apparent visual movement.

Since stimulus-conditions of apparent movement may be loaded by familiarity, expectancy, and meaning,[4] a test free from bias may be difficult to construct.[5] Furthermore, such a test may yield scores having correlated means and variances which would mean that they are not subject to parametric analysis.[6] If these objections can be met, we believe that the perception of apparent visual movement would provide a useful index for research in aging and functional efficiency.

This study is the result of an attempt to develop stimulus-conditions that will be free from these objections. We believe that this task may be accomplished (1) by describing the physical system used to control the stimulus; (2) by describing the distribution of responses in a population; and (3) by showing that the stimulus-situation provides a valid and reliable index of psychopathology or functional efficiency.

Apparatus. Optimal apparent movement (Beta motion) was produced by presenting two lights, 180° out of phase, with no pause (p) between them.[7] Two

* Received for publication August 23, 1959.

[1] R. T. Saucer and H. L. Deabler, The perception of apparent motion by organics and schizophrenics, *J. consult. Psychol.*, 20, 1956, 385-389; G. A. Talland, Perception of apparent movement in Korsakoff's psychosis, *J. Pers.*, 26, 1958, 337-348.

[2] Saucer, A further study of the perception of apparent motion by schizophrenics, *J. consult. Psychol.*, 22, 1958, 256-258; Saucer and Deabler, *op. cit.*, 388.

[3] Saucer, Chlorpromazine and apparent motion perception by schizophrenics, *J. consult. Psychol.*, 23, 1959, 134-136.

[4] E. E. Jones and J. S. Bruner, Expectancy in apparent visual movement, *Brit. J. Psychol.*, 45, 1954, 157-165.

[5] Saucer, *op. cit.*, 1958, 256.

[6] Saucer, *op. cit.*, 1959, 135.

[7] Saucer and Deabler, *op. cit.*, 385.

circular test-patches, which subtended approximately 1° of the visual angle, gave rise, when excited at serial rates ranging from 2 to 15 ~, to optimal apparent movement in almost all normal *Ss*. Since dark-adaptation is a critical variable, the use of saturated red lights of approximately 6700 Å tends to stabilize the cumulative effects of light-adaptation when the method of limits is used.

The stimulus-condition used in this experiment consisted of four neon indicator lights (supplied by Allied Radio Corporation, catalog number 52E504, used with the red plastic caps unaltered) mounted at the corners of a 6° 30′ square about the center of a 12 × 12 in. black Masonite board. The board was then inserted in slots cut in the sides of an optic bench in the form of a plywood box 14 × 14 in. square and 42 in. long, painted inside and out with non-reflecting paint and having a viewing port at one end. A viewing hood was used to hold *S*'s head in position, 36 in. from the Masonite board.

The four lights were so wired that apparent motion could be seen as a pair of lights moving up and down together or a pair of lights moving from side to side together. The lights at the upper right hand corner and the lower left hand corner of the square were wired out of phase. A DPDT switch allowed *E* to connect the upper light in parallel with the light across the square or the light below, while the light in the lower left hand corner was connected in parallel with the light across the square or with the light above. Thus, the two lights at the top or right side are in parallel but out of phase with the two lights at the bottom or the left side. This arrangement allowed *E* to ask for direction as well as quality of movement as a check on premature or anticipatory responses.

In general, the lights may be excited by any square wave generator covering the range of 2-20 ~ arranged to gate two tubes capable of carrying 5MA at 100-v. DC. Two gate tubes are required to maintain the necessary 180° phase relationship between the two lights. A suitable instrument, requiring a minimum range of skills and equipment for construction, has been described by Brown and Saucer.[8]

In this study, the lights were driven by a special multivibrator square wave generator covering the range 1-1000 ~ with fixed off-on ratios in steps between 1 : 1 and 10 : 1, producing a square wave with negligible ramp-off (less than 1%) and a rise-time of less than 1 μsec. over a current range of 0-50 m. amp. Excitation current for the neon lights was 1.25 m. amp. per light, measured by a 0-1 milliameter, with a 1 : 1 light-dark ratio.

In addition to the gate-pulse for the light-source, the instrument provided a trigger pulse to initiate a timed sweep from a Tektronix 162 Waveform Generator. The current waveform was then displayed on the calibrated 3-in. screen of a Tektronix 360 Cathode Ray Indicator and the pulse-time read directly off the calibrated screen.

The generator, the Tektronic system, and a small intercom amplifier were concealed in a plywood operating cubicle outside the direct view of *S*. A small red light provided illumination for recording data and for checking instrument settings. The entire room was light-proofed and air-conditioned to 70° F. The only auditory stimulation in the room was a constant, minimal hum from the air-conditioner motor.

Subjects. The Ss (77 men) were divided into three groups and ordered on a

---

[8] C. C. Brown and R. T. Saucer, *Electronic Instrumentation for the Behavioral Sciences*, 1958, 75-76.

rough continuum ranging from (1) non-institutionalized, self-supporting, young men, through (2) institutionalized, partially employable, older men, to (3) institutionalized, infirm, elderly men. The sample was drawn from the employee-member population of a Veterans Administration domiciliary. Group 1, young employees (N 27), ranged in age from 21 to 36 yr., were gainfully employed in non-professional positions, and were functioning satisfactorily in a community setting. Group 2, 'member' group (N 27), ranged in age from 42 to 73 yr. They required institutional living because of disabilities other than those involving the central nervous system. Their condition permitted them to engage in 8-hr. daily, paid, clerical duties. Group 3 (N 23) ranged in age from 42 to 71 yr. This group, hereafter referred to as the 'hemiplegic' group, had histories of cerebral vascular accidents with resid-

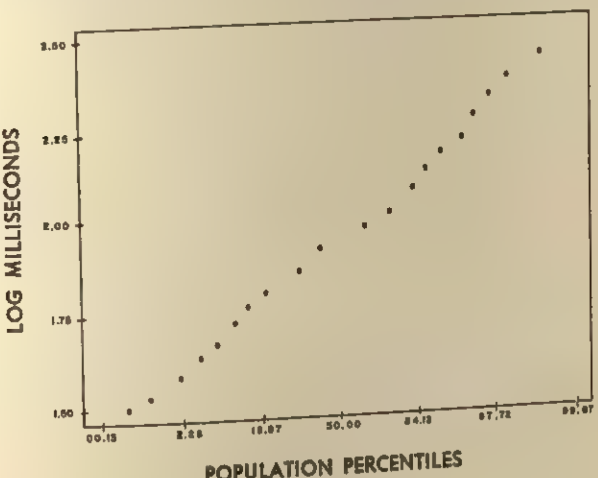


Fig. 1

ual hemiplegia. Each was ambulatory and could perform all activities of personal hygiene and self-care in an institutional setting. This trichotomy represents no known clinical entities but we believed that it would provide a wide range of global functional efficiency where the extremes are marked by youthful health and by age with known organic damage.

Procedure. Ss were brought into the test-room and asked if they would consent to serve in the experiment which was described as dealing with the effect of 'strokes' on visual perception. They were then given 3 min. of dark-adaptation and shown the stimulus-configuration in a low-rate condition to insure their being able to perceive and identify the movement. None of the Ss of Groups 1 and 2 failed to perceive apparent movement, but 12 of 40 hemiplegics within the Institution failed to report the perception even after it was described to them. Of the 28 who perceived motion, 5 were not available for retesting, hence Group 3 was reduced to an N of 23.

After identification of the phenomena, the Ss were given 20 alternate ascending and descending runs to threshold of perception of apparent movement. This order was further compounded by alternation in the horizontal and vertical dimensions.

When *S* reported the appearance or disappearance of movement, the pulse-time was read from the Tektronix Synchroscope.

Ss were retested at random, between 1 and 30 days following initial testing. The same procedure was followed for the retest-trials. The actual testing was randomized between the two *Es*, each testing approximately half the *Ss*, but in no pre-arranged order.

Results. As a first step in the analysis of the data, the cumulative frequency of the distribution of the raw scores was plotted. This was a logarithmic distribution. If the cumulative frequency is plotted against the $\log_{10}$ of the raw scores, the essentially normal distribution shown in Fig. 1 is given. We assumed, therefore, that the distribution of apparent thresholds of movement, in a sample population, is 'log-normal' for pulse-time. Parametric statistics may then be applied to data gathered from the stimulus-conditions described.

The hypothesis that these groups might represent a nearly pure age-

TABLE I

ANALYSIS OF VARIANCE BY GROUPS AND BY DAYS

Source	Sum of squares	*df.*	Mean square	*F*
Groups	.1518	2	.0759	6.78
Days	.0294	1	.0294	
Day × Group	.0004	2	.0002	
Remainder (Within)	1.6557	148	.0112	
Total	1.8374	153		

continuum was tested by a product-moment correlation between age and log score. The resulting correlation is 0.410, significantly different from zero for 77 *Ss*, but not exceptionally high. From the point of view of a variance-model, age contributes only 16% of the total variance.

The hypothesis that the grouping by level of functional efficiency would give rise to significant differences in log thresholds was tested by analysis of variance. The results are given in Table I.

There was no significant day-by-group interaction. The *F*-ratio of 6.78 for the group within comparison is significant beyond the 0.01 level of confidence, but probably not beyond the 0.001 level. It is believed that the null hypothesis of no true difference between groups may be rejected at an adequate level of confidence.

Since there was no day-by-group interaction and no significant difference between group means, it was assumed that the two test-periods represented equivalent test-situations. A test-retest correlation of 0.67 with a reliability of 0.80 indicates a reasonable degree of reliability for the measure.

The means and their associated standard deviations are presented in Table II.

Discussion. The techniques of the present study were designed, in part, to provide a set of normally distributed test-scores for use in parametric analysis of data. We believe that this objective has been attained; that with minimal requirements for instrumentation, a log-normal set of test-scores may be obtained.

At the same time, a set of normal raw scores may be obtained by measuring the height of a large sample without advancing our knowledge of the psychological processes of the individuals measured. The criterion of psychological utility must be satisfied.

We believe, furthermore, that the results of this study are useful. In a first approximation to a study of apparent movement and hemiplegia, we

TABLE II
MEANS AND STANDARD DEVIATIONS

Ss		*M*	*SD*
Employees	Test	1.998	.271
	Retest	1.975	.289
Member controls	Test	2.052	.129
	Retest	2.025	.103
Hemiplegics	Test	2.079	.200
	Retest	2.045	.138

find that 30% of the ambulatory, non-psychotic hemiplegics tested failed to perceive apparent motion under any set of conditions.

The perception of apparent motion is affected by organic damage, by schizophrenia and by chemotherapy, and shows a significant correlation with age. The results of this experiment may aid in finding an acceptable frame of reference for these diverse findings.

We believe also that the simplest frame of reference is that of our experimental criterion in this study. When groups of *Ss* are classified according to a very crude scale of behavioral efficiency, analysis of variance indicates significant differences in the ability to perceive apparent movement.

It has been stated that the effect of schizophrenia on the perception of apparent motion could not be separated from the effects of organic damage.[9] Age evidently produces an effect equivalent to either schizophrenia or organic damage.

All three of these effects may be regarded as being due to a generalized psychological deficit, the perception of apparent motion being correlated with the functional efficiency of the individual.

[9] Saucer and Deabler, *op. cit.*, 388.

SECONDARY REINFORCEMENT IN NEW LEARNING-SITUATIONS

By CLAUDE B. ELAM, Texas Christian University, and
D. W. TYLER, Louisiana State University

The principle of secondary reinforcement is based upon two empirical generalizations: first, that resistance to extinction is increased by stimuli which have been previously associated with primary reinforcement; and, secondly, that such stimuli may serve as reinforcing agents in a new learning-situation. The validity of the first generalization has recently been challenged by the finding that stimuli associated with non-reinforcement may, under certain conditions, produce greater resistance to extinction than do stimuli associated with reinforcement.[1] The present study presents an analogous challenge to the second generalization. A learning-situation is described in which stimuli associated with non-reinforcement were more effective in producing and maintaining a learned response than stimuli associated with primary reinforcement.

Subjects. Twenty male albino rats of the Wistar strain were tested in this study. They ranged in age from 120–160 days during the period in which the data were collected.

Apparatus. The apparatus used was an elevated *T*-maze with detachable arms. Removal of either arm formed an elevated *L*-shaped runway. The three sections (approach-stem and two arms) were painted mid-gray and were 18 in. high, 2 in. wide, and 24 in. long.

Two separate goal-boxes, 13 × 13 × 8 in., were also used. Except for the front panels, the exteriors of these boxes were identical. Their sides and backs were painted gray to match the runway, and their tops were covered with sheets of opal glass. One of the boxes had a white interior, a white front panel, and a floor covered with hardware-cloth. In addition, the white box had a guillotine door (also white) in the front panel. The other box had a black interior, a black front panel, and a smooth floor. Throughout all portions of the experiment, the goal-boxes were so situated at the end of the maze that the entrance always faced away from the starting point. The black and white panels were, therefore, not visible at the turning point.

Training. In the training stage, the *Ss* were given 20 spaced trials (15-min. minimal

* Accepted for publication September 1, 1958.

[1] M. E. Bitterman, W. E. Feddersen, and D. W. Tyler, Secondary reinforcement and the discrimination hypothesis, this JOURNAL, 66, 1953, 456-464; C. B. Elam, D. W. Tyler, and M. E. Bitterman, A further study of secondary reinforcement and the discrimination hypothesis, *J. comp. physiol. Psychol.*, 47, 1954, 381-384; D. W. Tyler, Extinction following partial reinforcement with control of stimulus-generalization and secondary reinforcement. this JOURNAL, 69, 1956, 359-368.

intertrial interval) per day for 10 days on the *L*-shaped runway. On half these trials, a right turn was required (left arm of the *T* removed) and on the remaining half a left turn was required (right arm of *T* removed). On half of the right-turning trials, *S*s encountered the white box at the end of the maze; the black box was presented on the remaining trials. This same procedure for presenting the boxes was followed for left-turning trials.

On half the trials with the white box present, the door in the front panel was open, and food (wet Purina Dog Chow) was present inside the box. On the remaining trials with the white present, the guillotine door was closed, and *S* was removed immediately following its approach to the closed door.

On all trials with the black box, the door of the box was open. On only one trial in 10, however, did *S* find food in the black box. Of the 200 trials in the training

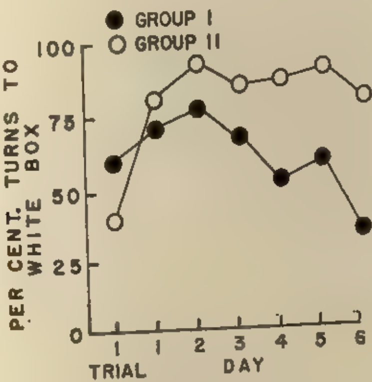


FIG. 1

stage, therefore, 50 were to food in the open white box, 50 were to the closed white box, 10 were to food in the open black box, and 90 were to the open but empty black box.

Test. At the start of testing, the *S*s were divided into two matched groups of 10 each on the basis of mean time required to traverse the runway from starting point to the turn during initial training. The *T*-maze, with the white box at the end of one arm and the black box at the end of the other arm, was used throughout the second stage. For any particular *S*, the position of the goal-boxes remained fixed throughout this portion of the experiment. The white box was on the left and the black box on the right for half the *S*s in each group, while the relative positions of the boxes were reversed for the remaining half of the *S*s in each group.

Each *S* received 10 spaced trials per day (a minimum of 15 min. again separated successive trials) for 6 days. The only difference between treatment of the two groups during the testing stage was that, for Group I, the door of the white box was open, while, for Group II, it always was closed. The black box was always open for both groups. Food was not present in either goal-box at any time during the testing stage.

Results. The percentage of turns to the white box for each group during the testing stage is shown in Fig. 1. Both groups developed an initial pref-

erence for turning to the white box. This preference for white was more pronounced and stable in Group II, which encountered the closed white box, than for Group I, which encountered the open, empty box. Over the 6-day period, the mean number of turns to white was 52.4 for Group II as compared with 36.6 for Group I. This difference is significant at the 1% level of confidence by Wilcoxon's test for paired replicates.[2]

The fact that the preference of Group II for the white box developed more rapidly and persisted longer is quite contrary to the expectation based upon the principle of secondary reinforcement. The *Ss* never were rewarded during initial training when the door to the white box was closed, whereas they were always rewarded when the door of this box was open. Under this training arrangement, approach to the closed white box could be expected to take on some generalized secondary reinforcement, but in no case could the strength of such generalized reinforcement be expected to be greater for those cues outside the box than for those cues inside. Accordingly, any difference in rate of learning which developed between the two groups should have favored Group I—the very reverse of the outcome obtained.

This difference between the two groups in preference for the white box appears to us to be interpretable in the same manner as the outcome of previous experiments by the authors and others in which the role of secondary reinforcement in extinction was studied. In these experiments on extinction, resistance to extinction of a response to a runway was significantly greater when the *Ss* were presented a goal-box of the same color as the one associated during acquisition with non-reinforcement than when they were presented with a box formerly associated with reinforcement.[3] In the present experiment, blocking *Ss* outside the white box on non-reinforced trials to that box was equivalent to presenting a box formerly associated with non-reinforcement. The *Ss* of Group II turned more consistently to the white box because training experiences and test-experiences were more congruous. The *Ss* of Group II simply could not discriminate readily that a significant change in procedure had been instituted during the testing portion of the study. Being blocked outside the white box on a portion of trials to this box was an integral part of training for all *Ss*, but being admitted to the white box without food (Group I) was quite contrary to the training procedure. The *inability*, then, of Group II to discriminate between the conditions of training and testing led to consistent selection of the white box by that group whereas the *ability* of Group I to discriminate led to a rapid turning away from the white box.

---

[2] Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures*, 1949, 1-16.

[3] Bitterman, Feddersen, and Tyler, *op. cit.*, 456-464; Elam, Tyler, and Bitterman, *op. cit.*, 381-384; Tyler, *op. cit.*, 359-368.

Reinforcing a small percentage of runs to the black box during initial training appears quite crucial insofar as obtaining the present results is concerned. In two previous unpublished studies by the authors, one with monkeys and the other with rats, in which a similar design was employed but without such reinforcement, no difference in rate of learning between two groups comparable to those in the present study appeared during test. The reinforcement procedure followed during initial training in the present study presumably resulted in a more sensitive test, since *Ss* had a more nearly equal tendency to turn in either direction.

Viewed in terms of resistance to extinction, the present results bear on the interpretation of the finding of Lambert and Solomon that resistance to extinction varies inversely as a function of distance from the goal at which blocking occurs.[4] This finding probably is restricted to cases in which no blocking occurs during initial training. When being blocked short of the goal forms a part of initial training, as in the present experiment, the relation between blocking and resistance to extinction may be quite other than that obtained by these investigators.

Summary. A study was conducted to determine whether stimuli immediately associated with primary reinforcement necessarily prove more effective in generating a new response than do stimuli associated with non-reinforcement. Twenty rats were given experience with the following conditions: (a) a closed white box, (b) an open white box containing food, (c) an open black box which contained food 10% of the time.

At the termination of this training the *Ss* were divided into two matched groups and run on a *T*-maze with the white box on one side and the black box on the other. No food was in either box during *T*-maze testing. The only difference between the groups was that the white box was always open for one group and closed for the other. It was found that both groups tended to approach the white box, but the group which found it closed learned somewhat more rapidly and maintained the response longer. This finding was discussed in relation to secondary reinforcement and the discriminative hypothesis.

[4] W. W. Lambert and R. L. Solomon, Extinction of a running response as a function of distance of block point from the goal, *J. comp. physiol. Psychol.*, 45, 1952, 269-279.

THE 'READY' SIGNAL IN EYELID-CONDITIONING

By WALLACE R. MCALLISTER and DOROTHY E. MCALLISTER, Syracuse University

In the study of eyelid-conditioning, one of the procedural decisions which must be made is whether to use a ready signal to indicate to the *S* the imminent presentation of the stimuli. Investigators have typically failed to make clear their reasons for the use, or lack of use, of such a signal. Their choice apparently has been based on pragmatic grounds. Little consideration seems to have been given to the possibility that the ready signal is a variable which might affect the level of conditioning.[1] It was the purpose of the present investigation to determine if such is the case.

Subjects. Thirty-two *S*s, 18 men and 14 women, from an introductory course in psychology participated in the experiment as partial fulfillment of the course requirement. Half of the men and half of the women, randomly assigned, were conditioned with a ready signal; the remainder, with no ready signal. Four additional women were used but were discarded for reasons discussed below.

Apparatus and procedure. The conditioned stimulus (*CS*) was a 1000 ~ tone at a sensation level of approximately 55 db. The unconditioned stimulus (*UCS*) was a puff of air of approximately 0.5 lb./sq. in. which was delivered to *S*'s right cornea through a nozzle 1 mm. in diameter. The tone was presented for 900 m.sec. and 800 m.sec. after its start the puff was turned on for 100 m.sec. Approximately 40 m.sec. were required for the puff to reach *S*'s eye. Thus, the interstimulus-interval was 840 m.sec.

The system used for recording the eyelid-movement and the procedure has been described in an earlier report.[2] The instructions read to all *S*s were the same except that those in the ready-signal group were informed that they would be told 'Ready' over the intercommunication system between the experimental rooms just before the presentation of the tone and puff. The ready signal actually was spoken either 3, 4, or 5 sec. prior to the administration of the stimuli, the interval varying irregularly from trial to trial. Each *S* was required to look throughout the experiment toward a rectangular, gray cardboard, 6-in. high and 16-in. wide affixed to a shield at about eye-level and approximately 24 in. in front of him. This procedure served to stabilize the base line of the recording. The remainder of the instructions were

* Received for publication October 7, 1958. The authors are indebted to Virginia Verhey for aid in conducting this experiment.

[1] An exception is a recent experiment, R. H. Dufort and G. A. Kimble, Ready signals and the effect of interpolated *UCS* presentations in eyelid conditioning, *J. exp. Psychol.*, 56, 1958, 1-7, published after the completion of the present investigation, in which this variable was studied.

[2] W. R. McAllister and D. E. McAllister, Effect of knowledge of conditioning upon eyelid conditioning. *J. exp. Psychol.*, 55, 1958, 579-583.

designed to produce a neutral set in *S* and withheld the nature of the experiment by informing him that the purpose was to study the reaction of the eye to different kinds of stimulation, i.e. tone and puff.

Prior to the conditioning trials, the *CS* was presented alone three times, at 10-sec. intervals to insure that *S* was following instructions and to determine whether *S*'s eyelid response was already conditioned to the tone. If *S* gave three conditioned responses (*CRs*), he was eliminated. Two women from the 'no-ready-signal' group were discarded for this reason. Following this, three presentations of the puff alone

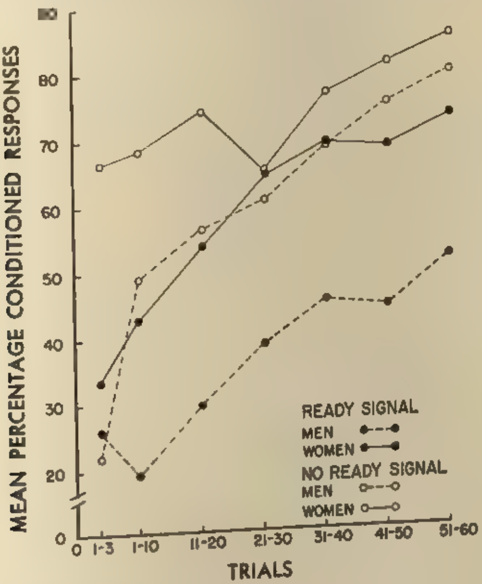


FIG. 1. PERCENTAGE OF CONDITIONED RESPONSES OF THE EYELID OVER BLOCKS OF TRIALS WITH AND WITHOUT A 'READY' SIGNAL

were given, at 10-sec. intervals, to acquaint *S* with its nature. After an interval of 30 sec., 60 conditioning trials were administered. The intertrial interval averaged 25 sec., varying in a fixed, irregular order of 20, 25, or 30 sec.

The measure of performance used was the percentage, in blocks of 10 trials, of anticipatory *CRs* defined as an eyeblink 1 mm. or more in amplitude, occurring between 250 and 840 m.sec. after the onset of the tone. A criterion for identifying voluntary conditioners was established beforehand in accordance with generally accepted procedures.[8] All *CRs* occurring between 250 and 460 m.sec. were considered to be voluntary responses. *Ss* with more than 50% of their *CRs* falling in this interval were classified as voluntary conditioners. Two women from the 'ready-signal' group met this criterion and were discarded.

[8] K. W. Spence and J. A. Taylor, Anxiety and strength of the *UCS* as determiners of the amount of eyelid conditioning, *J. exp. Psychol.*, 42, 1951, 184; Spence, Learning and performance in eyelid conditioning as a function of intensity of the *UCS*, *ibid.*, 45, 1953, 59.

Results. In Fig. 1 is presented the mean percentage of *CRs* as a function of blocks of trials for the four groups. From inspection of the figure it is apparent that, within each sex, the group without the ready signal responded more often than the group with it. It is also clearly indicated that, within each of the ready-signal conditions, the women responded more frequently than the men. The performance of the women who served without a ready signal seems to be unusually high at the beginning of practice. It is plausible that this is an artificial effect. In another study conducted subsequently, 15 women *Ss* served under similar conditions. Their performance-curve was quite comparable to that of the present group except for the initial trials. Over the first three trials their mean was 17.7%; over the first 10 trials, 53.7%.

TABLE I
SUMMARY OF ANALYSIS OF VARIANCE OF PERCENTAGE RESPONSES (CR) DURING EYELID-CONDITIONING

	Source	Mean	*df.*	*F*
			(31)	
Between *Ss*	Ready signal (RS):		1	5.66†
	With	48.8		
	Without	69.7		
	Sex (S):		1	3.77*
	Women	68.9		
	Men	51.8		
	RS×S		1	<1
	Error (b)		28	
			(160)	
Within *Ss*	Trials (T)		5	14.65‡
	RS×T		5	<1
	S×T		5	<1
	RS×S×T		5	<1
	Error (w)		140	
Total			191	

* $0.05 < P < 0.10$.
† $P < 0.025$.
‡ $P < 0.01$.

The differences in trends of the performance curves were evaluated by means of Lindquist's Type III analysis of variance.[4] A summary of the analysis and the means of the between-*Ss* main effects are presented in Table I. As may be seen in this Table, two of the factors, ready signal and trials, were significant beyond the 5% level. The third factor, sex, ap-

[4] E. F. Lindquist, *Design and Analysis of Experiments*, 1953, 281-284. The assumption of homogeneity of variance of the independent estimates of error (b) was tested with Bartlett's test and found to be tenable. A similar test was made with the independent estimates of error (w) with the same result, although, in this case, it is not a direct test of assumption.

proached significance. None of the interactions was significant. The results of this analysis together with an inspection of the figure, indicate that the use of the ready signal decreased the level of performance, and that learning took place in all groups at about the same rate.

Discussion. Although it is clear from this investigation that conditioned eyelid-responses occur less frequently when the procedure involves the use of a ready signal, the data provide no evidence to support a particular explanation of the results.[5] Two interpretations may, however, be suggested for consideration: (a) set; and (b) external inhibition.

Interpretation in terms of set is based upon the observation that the *Ss* frequently report a tendency to blink to the ready signal. Since the instructions explicitly requested *S* to keep his eyes oriented toward the cardboard (for recording purposes), the tendency to blink might be inhibited in order to follow the instructions. If this inhibition of the eyelid-response persisted until the conditioning trial was presented, it might be expected that the *CR* would be inhibited. Such inhibition would not have an opportunity to develop in the group without the ready signal.

The explanation in terms of external inhibition stems from the findings of Pavlov.[6] He reported that an extra stimulus introduced just prior to a *CS* would inhibit a *CR* by eliciting another, competing response. If the ready signal can be considered to be an external inhibitor, this explanation could account for the data. In this experiment, however, it is difficult to identify the competing response. The argument may be advanced that the competing response is that of not blinking, mediated by self instructions elicited by the ready signal. If such is the case, however, the two explanations may be the same.

Summary. Eighteen men and 14 women were given 60 trials in an eyelid-conditioning experiment, half of each sex with a ready signal given prior to the paired *CS-UCS* presentations and half without the ready signal. The women responded more often than the men, although the difference only approached significance. The *Ss*, regardless of sex, responded more frequently when there was no ready signal.

[5] The results obtained by Dufort and Kimble (*op. cit.*) furnish additional evidence that the ready signal interferes with eyelid-conditioning.

[6] I. P. Pavlov, *Conditioned Reflexes*, G. V. Anrep, trans., 1927, 44-47.

PERFORMANCE AS A FUNCTION OF DRIVE, REWARD, AND HABIT-STRENGTH

By John P. Seward and Donald M. Procter,
University of California, Los Angeles

Evidence has been presented elsewhere for interaction between strength of drive (D) and amount of reward (K) in their effect on running speed.[1] These results favored Hull's multiplicative relation[2] rather than Spence's additive one,[3] at least for the values of D and K employed.

Open to question, too, are the 'rules' for combining habit-strength (H) with the motivational variables to produce reaction-potential (E). To support Hull's conjecture that $E = f(D \times H)$, Spence has gathered impressive data, ranging from simple conditioning in rats and humans to the learning of paired-associates.[4] All this evidence reduces to the principle that a difference in performance produced by practice is enhanced by an increased level of drive, *i.e.* to a form of interaction between D and H. Similarly, Spence has cited divergent curves of practice with different amounts or delays of reward as possible evidence that $E = f(K \times H)$.[5] At the same time, he has noted conflicting evidence and emphasized the uncertainty of these formulas.

A previous attempt failed to clarify the role of H, presumably because the running response we used was already too dominant to be strengthened by practice. In the present experiment, we chose a response, bar-pressing, not already in S's repertory. A $2 \times 2 \times 2$ factorial design was arranged: half the Ss were tested under strong drive (D), half under weak (d); half of each group received a pellet as reward (K), the other half none (k); and half of each subgroup had been previously practiced in the response (H), while the other half had not (h). There were thus eight treatment-combinations, as follows: *DKH, DKh, DkH, Dkh, dKH, dKh, dkH,*

* Received for publication September 25, 1959. The work was aided by a grant from the Faculty Research Committee, University of California.

[1] J. P. Seward, R. A. Shea, and David Elkind, Evidence for the interaction of drive and reward, this Journal, 71, 1958, 404-407.

[2] C. L. Hull, *A Behavior System*, 1952, 7.

[3] K. W. Spence, *Behavior Theory and Conditioning*, 1956, 197 f., 215.

[4] Spence, *op. cit.*, 83 ff., 104 ff., 213 f., 221 ff.; A theory of emotionally based drive (D) and its relation to performance in simple learning situations, *Amer. Psychologist*, 13, 1958, 131-141.

[5] Spence, *op. cit.*, 1956, 196.

dkh. This design enabled us to compare the effects of each variable at different levels of the other two, thereby locating significant interactions.

Subjects. Female albino rats, 3–5½ mo. old, were randomly assigned to the eight groups, with the restriction that the groups were to remain as nearly as possible equal in size. Eighty-three *S*s in three replications started the experiment. Thirty-one were eliminated for failure to adjust to the apparatus during preparatory stages—14 in adaptation, 13 in habit-training, 4 in reward-training.[6] Fifty-two *S*s were tested, six or seven in each group.

Apparatus. A wooden box 28 × 8 × 7½ in. consisted of two approximately equal compartments, a black goal-box and an unpainted response-area. The goal-box contained a glass furniture-coaster, resting on a white card to make it more visible. Projecting from the unpainted half a wooden start-box, 7½ × 4 × 8 in., also unpainted, connected with the response-area through two sliding doors, one opaque, the other transparent. Raising the transparent door automatically started a 1-r.p.s. timer.

The two compartments were joined by a passage 4 × 4 in., in the middle of which a swinging door was held shut by a concealed relay. Made of Plexiglas, the door was painted black before the third replication to remove reflections. A retractable bar of ¼-in. aluminum tubing was pivoted outside the box. Its free end projected into the response-area through one wall 1¼ in. from the entrance to the passage and 1¼ in. from the floor. In test-trials, the bar protruded 1¾ in. from the wall; in the course of habit-training it was lengthened by 1 in. but was shortened again on the final day. Movement of the bar from ⅛ to ⅜ in. in any direction, requiring pressure of not more than 5 gm., released the door to the goal-box and stopped the timer. A push-button enabled *E* to produce the same results.

The entire apparatus was covered with black window-screen and lighted by two shaded, 7-w. Nite-lites, one over the start-box, the other over the middle passage. By conducting the experiment in a darkened room, we aimed to reduce visual distraction and to provide a decreasing gradient of illumination from start to goal. (Increasing the steepness of this gradient for two days during preliminary training in the second replication did not noticeably affect performance.)

Procedure. The *S*s were housed four in a cage with water always available. They were fed mash once a day for a period progressively reduced from 1 to ½ hr. All preliminary training was done before the daily feeding.

(1) Adaptation. In Replications 1 and 3, this period lasted six days. On Day 1, *S*s explored the apparatus in groups for 20 min. with the bar retracted and all doors open. On Day 2, each *S* explored alone for 5 min. On Days 3 and 5, all *S*s had two trials in rotation from the closed start-box with the swinging door open. On Days 4 and 6, they had three rotated trials in which *E* released the door to the goal-box when *S* put its head in the passage.[7]

---

[6] Sharp clicks of the door-opening and timing devices apparently caused most of the trouble. Since *S*s in Group *H* that refused to press the bar had to be replaced, a possible selective factor must be recognized. Its only obvious influence on the results would be to increase the apparent effect of habit-training.

[7] In Replication 2, the last three days of adaptation were omitted by mistake; Group *h*, however, received Day-4 treatment on the third and fifth days of habit-training.

(2) Habit-training. This stage was designed to give differential practice in opening the door by pressing the bar. Three trials a day were rotated within two groups comprising all *Ss*. For Group *H*, the procedure was like that of Groups *DkH* and *Dkh* in test-trials. The coaster in the goal-box was empty, the door closed, the bar accessible. *E* put *S* in the start-box; when *S* faced front, *E* raised the opaque door, counted three, and raised the transparent door. If *S* had not pressed the bar in 2 min., *E* waited till its head was close to it, then released the door. *S* was removed 30 sec. after entering the goal-box. For Group *h*, the procedure was the same except that the bar was withdrawn and the door to the goal-box already open. Thirty-one *Ss* in both groups followed these routines for 10 days; 21 *Ss*, for 8 days.

(3) Reward-training. Three rotated trials a day for four days were given, with door open and bar retracted. Group *k* found the coaster empty; Group *K* found it baited with a ½-gm. soft pellet. Both groups were allowed 30 sec. to explore the goal-box and eat the pellet, respectively.

(4) Test-trials. These continued for eight days. All four *D*-groups had their trials just before feeding; *d*-groups had theirs just after. As before, *K*-groups were

TABLE I

MEAN SCORES (100× RECIPROCAL OF LATENCY) FOR EIGHT DAYS OF TESTING

Habit-strength	Drive-reward conditions				Mean
	DK	*Dk*	*dK*	*dk*	
H	64.0	26.0	16.8	29.5	34.1
h	38.9	20.2	10.7	15.9	21.4
Mean	51.4	23.1	13.8	22.7	27.8

given a ½-gm. pellet in the coaster; for *k*-groups it was empty. All *Ss* followed the same procedure as Group *H* in habit-training. Three trials a day were rotated within the *D*- and *d*-divisions; intertrial intervals varied with trial-time, but the median interval was probably about 15 min.

Scoring. Raw scores were the latent times from raising the start-door to pressing the bar. These were transformed to reciprocals × 100, enabling us to assign a score of 1 to all times over 120 sec. and to use parametric tests. To equalize the groups for factorial treatment, four *Ss* were eliminated: two by random numbers, one that did not learn the response until the eighth day, and one that did not learn it at all.[8] Only one score per day, the median, was used for each *S*.

Results. Table I shows the mean scores of the eight groups averaged over all test-days. Note that *H*-scores are consistently higher than *h*-scores, and that the *D*-groups, with one exception, score above the corresponding *d*-groups; but, whereas *K* exceeds *k* when drive is strong, the opposite is true when drive is weak, and this reversal occurs at both levels of habit-strength.

Analysis of variance showed the three main effects, *D*, *K*, and *H*, sig-

[8] The last two omissions helped to compensate for the non-learners eliminated from Group *H* in habit-training.

nificant at the 0.1, 5, and 0.5% levels, respectively. The interaction between D and K was also highly dependable ($p = 0.001$), but H failed to interact significantly with D or K or $D \times K$. Further analysis of simple effects revealed that D was effective in the K-groups ($p < 0.001$), but not in the k-groups. Conversely, the difference between DK and Dk was significant at the 0.1% level, but the opposing difference between dK and dk fell short of the 5% criterion.

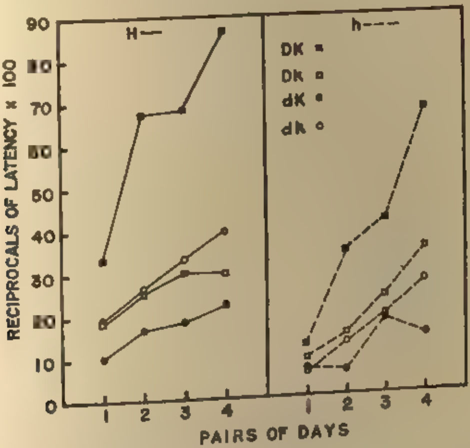


FIG. 1. MEAN SCORES FOR SUCCESSIVE PAIRS OF TESTING-DAYS

Fig. 1 shows mean scores averaged by successive pairs of days. A general upward trend is clear, with the curves for DK diverging steeply from the rest. The right-hand set (h) seems lower and more compact than the left (H), but its slopes and pattern are quite similar.

An analysis of variance was performed using order of trials (O) as an intra-S variable with 3 df. The order-effect, significant beyond the 0.1% level, interacted with D ($p = 0.005$) and with $D \times K$ ($p = 0.001$), but not with K or H or any interaction involving H. On further analysis, the O-effect proved significant ($p = 0.001$) at both levels of D. The $O \times D$ interaction, however, was significant only in the K-groups; conversely, the $O \times K$ interaction was significant only in the D-groups. Both these effects reached the 0.1% level.

Since practice on the test would be expected to diminish the difference between the H- and h-groups, we were interested in early scores, such as those represented by the initial points in Fig. 1. Analysis of the combined data for the first two days, however, gave results quite similar to those for the entire test. The effects of D and $D \times K$ were significant ($p =$

0.05), but K was not. Although the H-effect was clear-cut ($p = 0.001$), it did not interact significantly with the other variables.

In view of the convincing H-effect, we went back to see how much learning was actually manifest during habit-training. Mean scores for Group H in four successive pairs of days were 6.7, 9.2, 14.8, and 13.1. With 3 *df.*, the effect of order fell between the 5 and 10% levels of confidence. A test of the first four days against the second four days gave an F significant at the 0.5% level.

Discussion. The results of varying drive and reward, as defined by our procedures, confirmed the interaction previously reported.[9] Part of the interaction in this case was due to the superior performance of dk to dK groups, suggesting a negative valence of food for satiated animals. Since this difference by itself was not statistically significant, however, it should not be given much weight.

Turning to habit-strength, we find the evidence conflicting. Assuming that our scores measured reaction-potential, which in turn was a product of $D \times H$, we predicted that the effect of habit-training would be greater when tested under a strong drive than under a weak one. Although the difference was in the expected direction, it was not significant. The same was true when the H-effects were compared with and without a food-reward. Finally, the triple interaction of $D \times K \times H$ implied by Hull's postulate[10] was suggested by the data, but not to a statistically acceptable degree. From these findings, it could be argued that the Hullian formula should read: $E = f\,[(D \times K) + H]$.

It is possible, however, that improvement during the test itself is a sign of increased habit-strength. If $E = f(D \times H)$, the curves at different levels of drive should diverge with continued practice. That they do so is suggested by Fig. 1 and confirmed by the interaction of $O \times D$. This part of the data agrees with the many pairs of diverging curves cited by Spence.[11]

It is possible that the apparent inconsistency is a statistical artifact of the experimental design. The effect of habit-training had to be tested against an error-term representing between-S variance, while practice during test-trials had only to exceed the much smaller variance within Ss. More precisely measured—*e.g.* by a larger N—interactions with H as well as with O might prove significant.

As the data stand, however, they compel us to consider another possi-

[9] Seward, Shea, and Elkind, *op. cit.*, 404-407.
[10] Hull, *op. cit.*, 7.
[11] Spence, *op. cit.*, 1958, 131-141.

bility. For six of the eight groups—all but *DkH* and *Dkh*—habit-training took place under different motivational conditions than the test. To use the habit in the context of *K*, with or without drive, or in the presence of *D*, with or without reward, may have required some further learning. The nature of such learning is problematical, but Woodworth's description of *sequence learning* may contain a clue:

> Typically the first thing learned about the sequence is S_2, the incentive present in the situation; the second thing learned is the signal role of S_1; and the third thing is some instrumental preparatory R_1 which differs with the incentive and with the opportunities afforded by the situation.[12]

In other words, a learned instrumental response is not simply a 'habit' waiting to be activated by a drive, but must first be fitted into a complete motivating sequence. Only when the habit is formed in a constant setting of drive and incentive does it become sensitized to quantitative variations in one or the other. Hull implied something of the sort when he wrote of his multiplicative formula for *sEr:*

> This equation assumes that *D, V, K,* and *J* remain constant during learning and response evocation. It is not general enough to cover cases where changes are made in these factors during the experiment.[13]

Further speculation is unwarranted until habit-strength has been varied over a wider range, or until the effect of a change in reinforcing conditions has been measured without loss of precision.

Summary. The problem was to test Hull's assumption that drive (D), reward (K), and habit-strength (H) combine by multiplication to produce reaction-potential (E). A factorial experiment was designed with two levels of each variable established prior to testing. The measure of response was the latency of pressing a bar that gave access to a goal-box. Eight groups of six rats each were tested with three trials a day for eight days. Results were treated by analysis of variance to reveal significant interactions.

Measured by mean scores for the entire test, all three variables proved effective in accelerating the response. The only significant interaction was $D \times K$.

Performance improved throughout the test. This effect interacted significantly with D and $D \times K$, but with nothing else.

The first result argues against the equation $E = f(D \times H)$; the second argues for it. Two interpretations are offered; one statistical, the other based on the notion of learning as the integration of a sequence.

[12] R. S. Woodworth, *Dynamics of Behavior,* 1958, 258.
[13] Hull, *Essentials of Behavior,* 1951, 58.

ENHANCED LEARNING OF A POSITION-HABIT WITH SECONDARY REINFORCEMENT FOR THE WRONG RESPONSE

By A. D. Baddeley, M.R.C. Applied Psychology Research Unit, Cambridge, England.

It has recently been suggested that while the concept of secondary reinforcement is very frequently used, direct supporting evidence is surprisingly sparse and equivocal.[1] Much of the relevant evidence is concerned with the observation that a response which is followed by a secondary reinforcing stimulus extinguishes more slowly than a response with no such reinforcement, but this phenomenon can be explained at least as effectively in terms of discrimination.[2] Direct evidence of the effect of secondary reinforcement on acquisition is less common, and such effects as have been demonstrated are rarely as marked as the theory would predict.[3] Once again the results are open to an interpretation in terms of discrimination.

The present study aims to examine a situation in which the sight of food would be expected to strengthen the wrong R if it acts as a simple secondary reinforcement, while aiding the correct R if it acts as a cue to the location of reward.

Method. Two groups of rats are set the problem of learning a simple positional response. In both groups, the correct response is rewarded by the delivery of food accompanied by the onset of a light. The experimental group can see the light and food in the correct goal-box even after making a wrong response, while the control animals see food and light only when they make a correct R. The crucial question is whether the sight of food from the incorrect goal-box will serve as information about the correct response, thus increasing the speed of learning, or whether it will interfere with learning by acting as a secondary reinforcement for the wrong response.

A similar situation has been used in a study of probability-learning in the paradise

---

* Received for publication October 20, 1959. This work was carried out at Princeton University. The author wishes to thank Dr. D. K. Candland for his assistance.

[1] J. L. Myers, Secondary reinforcement: A review of recent experimentation, *Psychol. Rev.*, 55, 1958, 284-301.

[2] M. E. Bitterman, W. E. Feddersen, and D. W. Tyler, Secondary reinforcement and the discrimination hypothesis, this Journal, 66, 1953, 456-464; L. B. Wyckoff, Toward a quantitative theory of secondary reinforcement, *Psychol. Rev.*, 66, 1959, 68-78.

[3] W. K. Estes, A study of motivating conditions necessary for secondary reinforcement, *J. exp. Psychol.*, 39, 1949, 306-310; L. B. Wyckoff, J. Sidowski, and D. Chambliss, An experimental study of the relationship between secondary reinforcing and cue effects of a stimulus, *J. comp. physiol. Psychol.*, 51, 1958, 103-109.

fish by Bush and Wilson, who found the group given secondary reinforcement after the wrong response learned more slowly than the control.[4] In the present experiment, consistent reinforcement is used and the work is done with rats.

Apparatus. The apparatus consisted of a Lashley jumping stand modified as follows: (1) Pathways, 12 in. long and 3 in. wide, led from the starting platform to the two doorways, permitting the rat to walk rather than jump. (2) Each doorway was covered by a black curtain instead of a stimulus-card. (3) Behind the curtains was a goal-box (18 × 18 in.) of unpainted wood with a hinged roof of half-inch wire mesh. The goal-box was divided by a wire-mesh partition. (4) A semi-automatic feeding-device above each compartment allowed *E* to deliver food pellets into a glass feeding-dish. Each dish was shielded with wire mesh to ensure that *S*'s bodily orientation when eating from the dish was opposite to that necessary for the correct turn at the choice-point; *i.e.*, if the left compartment was correct, then the rat was forced to face right when actually eating. (5) In the corner of each compartment was a small 5-w. lamp, covered with a layer of white paint to give a soft but readily visible light. This lamp was switched on whenever food was delivered.

Procedure. Two groups of 10 hooded male rats from the Princeton University colony were used. Their age was approximately 120 days at the time of the experiment; they had been tamed by daily handling, and had previously been used on an elevated runway in a latent learning study. Following two weeks of deprivation which reduced them to 80% of their normal weight, the experiment began.

Pretraining was as follows. The goal-box and curtains were removed, and *S* was given 5 min. of exploration. On the next two days, a dish containing mash was placed in the center of the goal-platform; *S* was allowed five runs, after each of which it was allowed to eat a little mash. On the following day this procedure was repeated except that *S* was fed pellets instead of mash. Six *Ss* with very marked positional preferences were forced to make the non-preferred response (by blocking the preferred doorway) to equate the number of reinforcements of each turn. The *Ss* were then divided into two groups of 10, equated for positional preferences. Two *Ss*, one from each group, died during the course of the study and were replaced by two others.

The following day, the experiment proper began. Each *S* was given five trials separated by 15–30 sec. As soon as *S* entered one compartment, the spring door was closed to prevent retracing. For the experimental group, the light was turned on and food was delivered in the correct compartment as soon as the door was closed, regardless of the response made, while for the control group only the correct response led to the occurrence of light and food. The *Ss* in both groups were left in the goal-box for 15 sec. The left response was made correct for both groups. All the *Ss* were given five trials a day to the criterion of five correct responses in one day. They were fed the 5–10 gm. of food necessary to keep them at 80% normal body weight at the end of testing each day, except for the last few *Ss* run, which were fed 20 min. after running.

Results. Fig. 1 indicates that Group I, given information when making the wrong response, learned more rapidly than Group II. *Ss* from the two groups were paired according to the number of left responses made in pretraining. A Wilcoxon matched-pairs signed-ranks test carried out on the

[4] R. R. Bush and T. R. Wilson, Two-choice behavior in paradise fish, *J. exp. Psychol.*, 51, 1956, 315-322.

total number of errors showed Group I to have made significantly fewer errors ($p < 0.05$, two tail).[5] A similar test showed the experimental *Ss* to have required fewer trials to reach the criterion than the control group ($p < 0.05$, two tail).

The superiority of Group I on the first day suggests that the two groups were not equal in the first place. In fact, however, Group I actually made more errors than Group II on the first training trial, the only one which would not be affected by the different learning conditions; three of nine *Ss* of Group I and four of nine of Group II turned correctly on Trial 1. Unfortunately, first trial data are not available for the two *Ss* replacing those that died. Each made only two correct responses on Day I, and the most unfavorable distribution possible for our hypothesis (that the *Ss* of Group I made a correct response on Trial 1, and those of Group II did not) would mean that the two groups started with an equal number of correct responses.

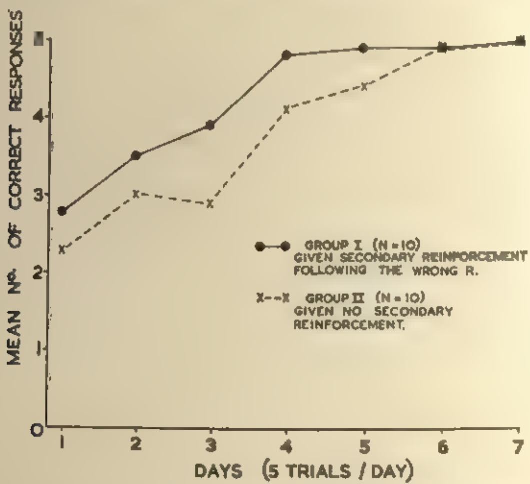


FIG. 1. MEAN NUMBER OF CORRECT RESPONSES PER ANIMAL PER DAY

Discussion. The main difficulty confronting a straightforward S-R reinforcement approach is to explain why the sight of food does not act as a secondary reinforcement for the wrong response and thereby hinder the learning of the correct response. There seem to be two possible approaches to this difficulty. First, *S* may be avoiding the incorrect compartment. This, however, would seem hard to justify either in terms of light-aversion, or frustration without seriously modifying secondary reinforcement theory. A second, more feasible explanation, however, might be as follows: The *Ss*

[5] S. Siegel, *Nonparametric Statistics*, 1956, 75-83.

might be expected to respond initially on the basis of the cue associated most closely with reinforcement. Bodily orientation while feeding might be such a cue. On making the correct response, all the *S*s were forced by the wire mesh shield round the feeding dish to orient to the *right* when feeding in the correct (*left*) goal box, but, the experimental group also would receive secondary reinforcement for left orientation when making the wrong response. If, however, the *S*s began by responding to the cue of bodily orientation, they would be expected to perform at a level of 50% or less until they abandoned it in favor of a more relevant cue. As the responses of the experimental group to this cue were partially reinforced, they would be expected to abandon it less readily than the controls. Since both groups show rapid improvement beyond the 50% level, an effect particularly marked in the experimental group, this explanation seems improbable.

The sight of food and light here appears to strengthen not the response it follows, but the alternative R. In common-sense terms, it appears to give the rat information about the location of reward. This interpretation may be expressible in mediational terms, but the question of why the sight of food does not act as a simple secondary reinforcement in this situation would still have to be answered. An approach to the phenomena of secondary reinforcement in terms of discrimination,[6] or of a cue-hypothesis,[7] while not necessarily predicting this result would have less difficulty in accommodating it, largely because the relationship between cue and response is not so rigidly specified as is that between S, R, and reinforcement in traditional S–R reinforcement-theory. To regard the sight of food and light in this study as a cue, or source of information about the situation is the least elaborate, and probably, in the present state of knowledge, the most fruitful approach.

Summary. Two groups of rats were presented with a positional problem on a modified Lashley jumping stand. The control group was trained using the non-correctional method. The experimental group was treated in the same way except that *S* was able to see the delivery of food and an associated light after making the wrong response. The experimental *S*s learned more rapidly and with fewer errors than the control group. It is suggested that the concept of secondary reinforcement has difficulty in accounting for this result.

[6] C. B. Elam, T. W. Tyler, and M. E. Bitterman, A further study of secondary reinforcement and the discrimination hypothesis. *J. comp. physiol. Psychol.*, 47. 1954, 381-384.

[7] L. B. Wyckoff, *op. cit.*, 68-78.

ON SIZE-PERCEPTION IN THE ABSENCE OF CUES FOR DISTANCE

By HANS WALLACH, Swarthmore College, and
VIRGIL V. MCKENNA, Carleton College

It is generally accepted that the veridical perception of size requires adequate cues for the distance of the object from *O*. When distance-cues are eliminated, attempts to match the size of similar objects at different distances produce equations such that the matched objects subtend equal visual angles, *i.e.* produce retinal images of equal size.[1] While the absence of veridical size-matches under these conditions is only to be expected, the regular occurrence of equated retinal images presents a problem. There are no other indications that there is immediate awareness of image-sizes or that one can react to them directly. That these image-matches occur may, however, be of great importance for the theory of size-perception.[2] In the absence of distance-cues, image-size alone may be able to determine perceived size. It is, therefore, essential to ask whether these matches really are significant facts in the perception of size. How persistent are they? Will they prevail when, without introducing conditions for veridical matches, the experimental situation is altered? We made such a change by arranging that, when two objects were to be compared, only one was presented devoid of cues for distance; the viewing conditions for the other provided the best cues for distance compatible with the necessary dark-room conditions.

Method. The standard stimulus-object was a luminous square 8 × 8 in., cut into a black cardboard that covered the face of a light-box. It was placed 108 in. from *O*'s eye and on a level with it. It was viewed monocularly through two circular apertures one behind the other and so aligned with the standard object that, looking through them, *O* could see the standard object but nothing else. The experimental room was completely dark. The comparison-objects, squares of white cardboard against a black background, were presented one at a time inside a large box into

* Received for publication August 14, 1959. This work was supported in part by the U. S. Naval Training Device Center under Contract N61339-511 with Swarthmore College.

[1] William Lichten and Susan Lurie, A new technique for the study of perceived size, this JOURNAL, 63, 1950, 280-282; E. L. Chalmers, Monocular and binocular cues in the perception of size and distance, this JOURNAL, 65, 1952, 415-423; A. H. Hastorf and K. S. Way, Apparent size with and without distance cues, *J. gen. Psychol.*, 47, 1952, 181-188.

[2] A. S. Gilinsky, The effect of attitude upon the perception of size, this JOURNAL, 68, 1955, 173-192.

which O could see with both eyes through an aperture of appropriate size. A lamp in the box illuminated its interior, but not enough light escaped through the aperture to affect the darkness of the experimental room. A headrest in front of the aperture served to fix the distance from the eyes to the comparison-object, which was placed against the far wall of the box. The distance between the eyes and the comparison-object was 27 in., one quarter that of the standard object. Thus, a square of 2 in. produced a retinal image as large as that of the standard square. The 25 comparison-objects ranged in size from ½ to 14 in. Twenty-six undergraduates served as Os.

The method of limits was used for the matches. Ascending and descending trials were alternated within the experiment and from S to S with regard to initial trials. Two ascending and two descending trials completed a matching experiment. The first ascending trial always started with the ½-in. square and the second with the ¾-in. one. The first descending trial began with the 14-in. square and the second with the 13-in. one. Each trial was continued until O had given smaller, larger, and equal judgments.

For 16 of the Os, a control experiment followed the darkroom-matches. The light in the room was turned on, the screens with the circular apertures were removed, and O was allowed to view the standard square binocularly. One ascending and one descending trial then were given under these constancy-conditions.

Results. Under reduction-conditions, only 7 of 26 Os made matches in which the retinal image sizes were approximately equated. The judgments of a majority, 17 Os, were strongly influenced by the size of the comparison-object. For ascending trials, which began with small objects, they judged small objects equal to the standard; for descending trials, which began with large objects, they judged large objects as equal. For this group of Os, the mean of their matches in ascending trials was 2.05 in., and in descending trials it was 9.6 in. Since the Os who showed this difference between ascending and descending trials were in the majority, there was also a large mean-difference between ascending and descending judgments for the whole group of 26 Os. This difference amounted to 5.2 in. Of the remaining two Os, one gave consistently large matches that averaged 7.8 in. and the other intermediate equality-judgments ranging from 4 to 5.5 in.

In the control experiment, where constancy-conditions prevailed, the difference between ascending and descending judgments was, of course, much smaller. The mean for ascending judgments was 7.94 in., and for descending judgments 9.33 in., close to the true size of the standard object, which was 8 in. square. The difference between ascending and descending judgments in the control experiment was reliably smaller ($p < 0.005$) than that obtained under reduction-conditions.

The experiment was repeated with a five-cornered nonsense-figure instead of the square and with a different group of Os. Essentially the same

result was obtained as before. Only 2 of 10 *O*s equated retinal-image sizes in ascending trials.

Conclusion. We found that in the absence of distance-cues for the standard object only a minority of *O*s produced matches in which retinal images were equated. The previous result, that such matches occur regularly when both the standard and the comparison-object are given without distance-cues, thus becomes a fact of minor importance. Evidence for this view is to be found in the nature of those matches which, in the case of a majority of our *O*s, replace the equation of retinal images. They make large matches when the comparison-objects, whose sizes can be readily perceived, are large, and small matches when comparison-objects are small. This inconsistent type of responding shows clearly that the standard object in reduced conditions does not elicit unambiguous size. That image-size in the absence of distance-cues has the power to determine size-perception directly is no longer a feasible assumption.

How, then, can one account for those matches in which image-sizes are equated? It does not seem particularly important to find an answer to this question, since it has been shown that these matches are readily replaced by an inconsistent way of responding, but we believe that the equation of image-sizes results from an implicit assumption of equal distance of the standard and the comparison-object.

APPARATUS

A NEW TEST OF SCOTOPIC SENSITIVITY

By Jo Ann Smith Kinney, Edward J. Sweeney, and Alma P. Ryan, U. S. Naval Medical Research Laboratory, New London, Connecticut

A new test of sensitivity for night vision has been developed and standardized at this Laboratory and it is now available in a compact, portable model. This paper outlines the principles upon which the test is based, summarizes the data obtained with an experimental model, presents the results found with this final version of the test, and indicates reasons for an optimistic prognosis for its use.

The test is based upon sampling the subject's (*S*'s) scotopic sensitivity at a number of retinal positions. Previous studies have shown several basic principles in this regard.[1] The sensitivity of the eye to dim light varies throughout the retina, being poorest at the fovea, increasing to a peak at 10 to 15° from the fovea, and then decreasing gradually to the extreme periphery. In addition, sensitivity is not uniform throughout the various quadrants, the nasal field, *i.e.* the temporal retina, giving the most sensitive measures. While averaging the data for a group of *Ss* yields the same map of retinal sensitivity, there are significant individual differences, not only in the over-all levels of sensitivity, but also in the profiles obtained.[2]

To evaluate *S*'s sensitivity properly it is thus necessary to test at various locations and to give him a score compiled from these various measures.

The test also utilizes stimuli of various sizes. Previous results have shown that small stimuli of different sizes yield the same type of data as do small stimuli of different brightnesses.[3] Normal psychophysical functions are found when size is used as the variable; in addition, identical maps of retinal sensitivity are obtained when the stimulus-variable is either size or brightness. This allows an extremely convenient measure of sensitivity, in terms of size, and also yields an internal check of the validity of *S*'s results.

Various methods of administering and scoring the test also have been investigated.

---

[1] J. A. Smith and F. L. Dimmick, The parameters of scotopic sensitivity: I. The effect of size, *Naval Medical Research Laboratory*, Report No. 174, 1951, 56-75; S. G. de Groot, J. M. Dodge, and J. A. Smith, Factors in night vision sensitivity: The effect of brightness, *ibid.*, Report No. 194, 1952, 1-17; Factors in night vision sensitivity: The interrelation of size, brightness, and location, *ibid.*, Report No. 234, 1953, 1-13.

[2] E. J. Sweeney, J. A. S. Kinney, and A. P. Ryan, Standardization of a scotopic sensitivity test, *ibid.*, Report No. 308, 1959, 1-8.

[3] Smith and Dimmick, *op.cit.*, 67, De Groot, Dodge, and Smith, *op.cit.*, No. 234, 12-16.

On the basis of these results, the present method of presenting two stimuli at a single trial and of using the total number correctly identified as the score was chosen.[4]

These then are the considerations and principles upon which the test is based. The shortcomings of many of the wartime tests of night vision, particularly with reference to reliability, are well-known and have been adequately summarized by Berry.[5] These difficulties were such that some investigators were led to doubt that practical differences in night vision existed in the normal population (that is, among those who were not night blind).

As knowledge about night vision was accumulated, many of the reasons for the failures became apparent. These include the large effects on sensitivity of previous exposure to light, not only with reference to the immediate hours before testing, but also long-term, seasonal effects of light-history. It has recently been shown, for example, that the amount of seasonal variation in sensitivity for a given individual is as great as the normal variation found between individuals at any given time.[6]

This increased knowledge of relevant variables, together with the proper care in controlling them, has revealed the existence of sizeable, reliable differences in night vision, and should result also in the valid testing of this ability.

The first model of the test, a large experimental version, was used on a sample of 108 enlisted men. The results were very encouraging. Highly significant differences in sensitivity were found between individuals. The test-retest correlations were in the low 0.80s, both when testing was done in one day and on consecutive days. Various measures were made to indicate the internal validity of the test, and all gave positive results. These measures were based upon the expectations that higher percentages would be scored for larger sizes of stimuli and that the previously established map of retinal sensitivity would be duplicated in the group-results.[7]

The present model was then constructed, using the same principles as the first one, but it was made considerably smaller and portable. This necessitated eliminating the large sphere that gave constant brightness at each position, but such constant brightness throughout the visual field was necessary only to test for the expected map of retinal sensitivity. The new model has been used to test another group of enlisted men to be certain that the positive results obtained with the experimental model would be duplicated.

Equipment and procedure. The new version of the test of sensitivity of night vision is shown in Fig. 1, as photographed from *S*'s point of view. It has been placed in a black booth which shields *S* from light and holds his chin-rest. *E* sits on the opposite side of the booth, screened from *S*; he presents the stimuli and records responses.

The light-source for the test, a miniature incandescent lamp, is surrounded by a short, white cylinder and enclosed within a box 14 in. square and 3 in. deep. Both

[4] Gilbert Fooks, E. J. Sweeney, and F. L. Dimmick, Pilot studies of a scotopic sensitivity test, *Naval Medical Research Laboratory*, Report No. 285, 1957, 1-7; Sweeney, Kinney, and Ryan, *op. cit.*, Report No. 308, 1959, 5-7.

[5] William Berry, Review of wartime studies of dark adaptation, night-vision tests, and related topics, *Armed Forces-NRC Vision Committee*, Dec. 1949, 1-96.

[6] Sweeney, Kinney, and Ryan, Seasonal changes in scotopic sensitivity, *J. opt. Soc. Amer.*, 50, 1960, 237-240.

[7] Sweeney, Kinney, and Ryan, *op. cit.*, Report No. 308, 3-4.

the cylinder and the front surface of the box, as Fig. 1 shows, have vertical and horizontal slits through which the stimuli shine. Between the cylinder and the front of the box is a large cogwheel made of sheet metal in which tiny holes of different sizes and positions have been drilled. The knob on the front turns the cogwheel to bring consecutive holes in front of the slits. In the center of the box is another miniature lamp with a deep red filter which serves as a fixation-point.

The source is powered by a 1½ v. battery connected to a Hunter Timer to give a 1-sec. exposure. The brightness of the white cylinder is not constant over the field, since it is not a perfectly diffusing sphere. Neutral filters and a color com-

FIG. 1. PHOTOGRAPH OF THE PORTABLE VERSION OF THE TEST OF NIGHT VISION

pensating filter have been placed over the slits; these produce a field of approximately equal brightness of 5.5 log μμL and 2042°K.[8]

Five sizes of stimuli are used in the test; at a viewing distance of 14.3 in., these subtend visual angles of 0.09, 0.124, 0.156, 0.186, and 0.220°. Twelve positions in the visual field are tested; these are located at 6, 12, and 20° above, below, to the right, and to the left of the central fixation-point. By the time the cogwheel has returned to its original position, each of the 5 sizes has been presented twice at each of the 12 positions.

Two spots are shown on each trial in different quadrants; in this way, the 120 stimuli are presented in 60 trials. *S* is told that he may see one or both of the spots and is instructed to report where he has seen the spots in terms of 'up,' 'down,' 'right,' or 'left.' *S* turns the knob between trials to reveal the next pair of stimuli. The entire test takes less than 15 min. excluding dark-adaptation.

[8] For the interrelations of the size, brightness, and color temperature of the source in determining thresholds, see De Groot, Dodge, and Smith, *op.cit.*, Report No. 234, 12-16; and J. S. Kinney, *J. opt. Soc. Am.*, 46, 1956, 1093-1094.

In the present evaluation of the test, 30 naval enlisted men served as *Ss*. Six of these were tested in February and the others in June and July. Their ages ranged from 17 to 33 yr.

S was dark-adapted for 20 min. before testing began. His left eye was occluded; his right eye used throughout the test. The entire 120 stimuli were presented, a rest-period given, and then the test was repeated to determine a test-retest correlation.

Results. Fig. 2 presents a frequency-distribution of the scores, in terms of the total number correctly identified, from the first session in which the

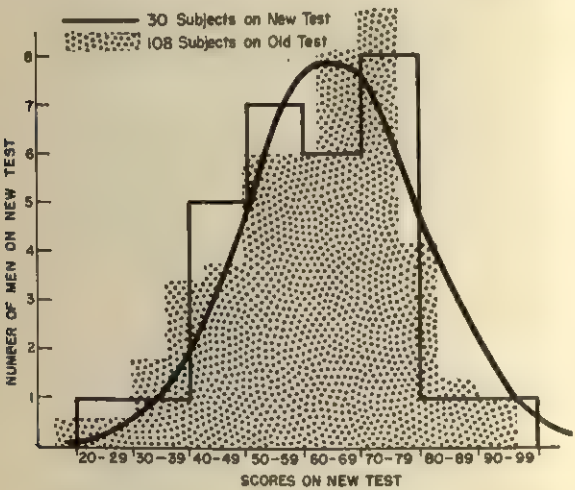


FIG. 2. DISTRIBUTION OF TEST-SCORES
(Total number correct.)

30 *Ss* were tested. The scores range from 28 to 95, showing large individual differences and a fairly normal distribution. The distribution for the second session was similar. The distribution of the larger group of *Ss* (108 men used in a previous study)[9] is superimposed by means of its own normal curve on the present distribution. The two distributions are very similar in shape, indicating that the new test is differentiating in the same manner as the old.

Table I gives the average data for the men in both sessions, together with the test-retest correlations. To compare seasonal effects the *Ss* were separated according to the seasons in which they served.

The average for the winter group is higher in both sessions than the summer group. It might be anticipated that this factor would inflate the test-retest correlations of the total group if, for example, there were a few

[9] Sweeney, Kinney, and Ryan, *op. cit.*, Report No. 308, 3.

extremely high scores in the winter group. To check this, the average difference between winter and summer scores was subtracted from the scores of each of the men in the winter group and the correlation for the total group was recomputed. This procedure resulted in a correlation of 0.894, as compared with 0.897 for the uncorrected scores; hence we may safely conclude that the high reliability of the test does not depend on seasonal variations.

As may also be seen in Table I, the scores obtained in the second session were somewhat higher than those in the first. The increase is in part due to practice, a result found also with the previous group of 108 *S*s, and in part to an increase in sensitivity due to an extra 15 min. of dark-adaptation by the time the second session was started. The latter, however, cannot affect the use of the test-results as long as conditions remain the same for all the *S*s.[10]

TABLE I

TOTAL NUMBER CORRECT RESPONSES WITH THE NEW VERSION OF THE NIGHT VISION TEST

Group	First session	Second session	Test-retest *r*
Summer (24*S*s)			.865
Mean	58.5	63.7	
SD	13.6	14.2	
Winter (6*S*s)			.962
Mean	66.7	69.8	
SD	17.8	19.9	
Total (30*S*s)			.897
Mean	60.1	64.9	
SD	14.9	15.7	

Since the test consists of stimuli of different sizes, the data for every *S* can be analyzed in terms of the number correctly seen at each size. This provides an internal check on the validity of *S*'s results, since he should see more of larger stimuli, and it also gives another means of evaluating the meaning of the individual differences.

In Fig. 3A, the average number correctly seen by the 30 *S*s at each size is plotted on probability paper for both the first and the second sessions. In Fig. 3B, the data for the most sensitive and the least sensitive *S*s are plotted for comparison. The close agreement between test and retest for the group can be compared with the large difference between the poorest and

[10] To determine whether the time allotted to adaptation could be conserved without altering the results, 24 men were tested after 10, 20, 30, and 40 min. of dark-adaptation. Correlations between the number correct at the different stages of dark-adaptation were as high between 20 and 30 min. as they were between 30 and 40, and indeed as high as the test-retest values taken after 30 min. of adaptation for each session.

best S. The poorest S required a stimulus more than 2-½ times as large as the best S to be seen 50% of the time and more than 1-½ times as large as the average S.

The results of this portable version of the test are in complete agreement with those of the larger, experimental model. The distribution of total scores is very similar; the test-retest correlation is slightly higher in this version; and individual differences are large and significant. Furthermore, checks of the pattern of retinal sensitivity revealed once again the

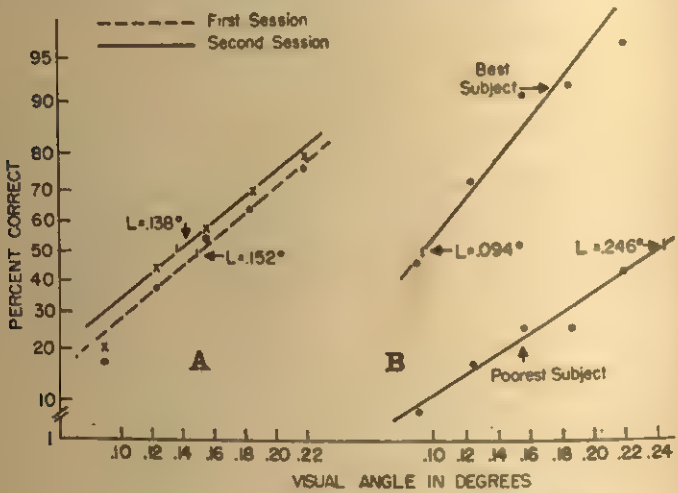


FIG. 3. VISUAL ANGLE AND PERCENTAGE CORRECT
A-average for 30 Ss in the first and second sessions;
B-sensitivity of the poorest and best Ss

same map for the group data with the 12° position and the temporal retina yielding greatest sensitivity.

Discussion. The problem of the external validity of the test is necessarily a complicated one. That the test measures what it is supposed to, *i.e.* sensitivity of night vision, has been shown by a variety of means, but this does not prove that it will be a predictor of performance at night. In evaluating performance, considerations of motivation, intelligence, and numerous other factors must be taken into account. There is no doubt that an alert, intelligent man with poor sensitivity may perform as well as a sleepy man with excellent sensitivity. On the other hand, it should be obvious that, given the same conditions in the two men, the one with superior sensitivity will give a better performance.

Thus an attempt to correlate test-scores obtained under dark-adaptation with performance at night must include controls for other factors relevant to performance. Such a project is a complicated one, but it could be undertaken for specific nocturnal tasks with the aim of weighting the various factors for their importance in prediction.

Without such corrections, one may raise the question as to what the differences in sensitivity between men mean in a practical situation. Since the results of the test can be given in terms of the size of visual angle that is seen a given percentage of the time, a direct comparison can be made.

The difference between the limens of the extremes yields a ratio of 2.6. This can be converted to the distance at which the same sized target would be visible. If, for example, the poorest *S* could see a particular light 50% of the time at a distance of 350 ft., it would be just as visible to the most sensitive *S* at 914 ft. The difference is, of course, considerable and could be of great importance provided both *S*s were equal in other respects.

Summary. A new test of scotopic sensitivity was developed. The test and its portable apparatus are easily reproduced and given. As a testing device, it employs principles which show promise of reliably differentiating *S*s according to their sensitivity in night vision. These include measures of sensitivity at numerous locations in the visual field, presenting two stimuli on each trial, and using different sized stimuli from which a total score is compiled.

This test has been used to examine 30 naval enlisted men. The scores show large individual differences and have a test-retest reliability of 0.89. The results were very similar to those obtained from 108 *S*s who were tested with a preliminary model. The questions of the meaning of the individual differences and of the external validity of the test are discussed.

NOTES AND DISCUSSIONS

PORTRAITS OF BRITISH PHILOSOPHERS AND SCIENTISTS

The following list may be of interest to psychologists. It contains references to portraits, usually photographs, of British psychologists and other persons such as anthropologists, philosophers, physiologists, and other scientists whose work will be known to students. British is defined loosely—Titchener and Anna Freud are included.

A somewhat similar international list was published some years ago by Bergman and Dallenbach.[1] Borrowings from this list are indicated by an asterisk (*).

AVELING, FRANCIS ARTHUR POWELL (1875-1941). (1) *Brit. J. Psychol.*, 32, 1941, facing 1. (2)* *Wittenberg Symposium*, 1928, 3.

ALEXANDER, SAMUEL (1859-1938). (1) *Brit. J. Psychol.*, 29, 1939, facing 319. (2) M. Münsterberg, *Hugo Münsterberg*, 1922, facing 26 (in a group of students with Münsterberg at Freiburg).

BAIN, ALEXANDER (1818-1903). (1) *Autobiography*, 1904, frontispiece. (2) *Ibid.*, 239. (3) *Ibid.*, 337. (4) *Ibid.*, 407. (5)* J. Sully, *My Life and Friends*, 1918, facing 86. (6) H. D. Traill and J. S. Mann, *Social England*, 6, 1904, 676.

BALLARD, PHILIP BOSWOOD (1865-1950). (1) P. B. Ballard, *Things I Cannot Forget*, 1937, frontispiece.

BARTLETT, FREDERIC CHARLES (1886——). (1) *A History of Psychology in Autobiography*, 3, 1936, XIV. (2) David Katz, *Psychologischer Atlas*, 1945, 89.

BELL, CHARLES (1774-1842). (1) *Harpers' Magazine*, 99, 1899, 512. (2) O. L. Bettmann, *A Pictorial History of Medicine*, 1956, 242. (3) T. H. Garrison, *An Introduction to the History of Medicine*, 2nd ed., 1917, 459. (4) Gordon Gordon-Taylor, *Sir Charles Bell, His Life and Times*, 1958, 16. (5) *Ibid.*, 152. (6)* F. Fearing, *Reflex Action*, 1930, facing 112.

BENTHAM, JEREMY (1748-1832). (1) C. W. Everett, *The Education of Jeremy Bentham*, 1931, frontispiece. (2) *Bicentenary Celebrations, Jeremy Bentham*, Tuesday 8th June, 1948, frontispiece. (3) *The Listener*, 60, 1958, (No. 1540), 503.

BERKELEY, GEORGE (1685-1752). (1) A. C. Fraser, *Life and Letters of George Berkeley*, 1871, frontispiece. (2) A. A. Luce, *The Life of George Berkeley, Bishop of Cloyne*, 1949, Frontispiece. (3) *Ibid.*, 149. (4) *Ibid.*, 180. (5) G. Simpson (ed.), *The Works of George Berkeley*, 1897, frontispiece.

BRAID, JAMES (c. 1795-1860). (1)* *Scient. Mo.*, 29, 1929, 157. (2) *Medical History* 1, 1957, 242.

BROWN, THOMAS (1778-1820). (1) Thomas Brown, *Lectures on the Philosophy of the Human Mind*, 1830, frontispiece.

BROWN, WILLIAM (1881-1952). (1) W. H. R. Rivers, *Psychology and Politics and other Essays*, 1923, frontispiece. (2) *Brit. Med. J.*, 1952, (May 24), 1136.

BURT, CYRIL LODOWIC (1883——). (1) *A History of Psychology in Autobiography*, 4, 1952, (2) *Occup. Psychol.*, 23, 1949, 9. (3) David Katz, *Psychologischer Atlas*, 1945, 90. (3) R. B. Cattell, *Your Mind and Mine*, 1934, 271.

---

[1] L. V. Bergman and K. M. Dallenbach, Portraits useful to the psychologist, this JOURNAL, 45, 1933, 165-171.

CATTELL, RAYMOND BERNARD (1905——). (1) *Human Affairs,* ed. by R. B. Cattell, John Cohen, and R. W. Travers, 1937, 140. (2) *Contemp. Psychol.,* 1959, 324.
COHEN, JOHN (1911——) (1) *Human Affairs,* ed. by R. B. Cattell, John Cohen and R. W. Travers, 1937, 6. (2) *Contemp. Psychol.,* 4, 1959, 228.
COMBE, GEORGE (1788-1858). (1) C. Gibbon, *Life of George Combe,* 1 and 2, 1878, frontispieces. (2) *Ibid.,* medallion on title page in both volumes.
CRAIK, KENNETH J. W. (1914-1945). (1) *Brit. J. Psychol.,* 36, 1946, 109.
CULPIN, MILLAIS (1874-1952). (1) *Occup. Psychol.,* 23, 1949, 140. (2) *Lancet,* 1952, (Sept. 27), 643.
DALTON, JOHN (1766-1844). (1) Julia Bell, *Colour Blindness,* The Treasury of Human Inheritance, 2, 1926, frontispiece Pt. 1. (2) W. C Henry, *Memoirs of the Life and Scientific Researches of John Dalton,* 1854, frontispiece. (3) *The Record of the Royal Society,* 3rd ed., 1912, 224, (Plate XIV).
DAWSON, SHEPHERD (1880-1935). (1) *Brit. J. Psychol.,* 26, 1935, 117.
DREVER, JAMES (Senior) (1873-1950). (1) *Brit. J. Psychol.,* 42, 1951, 311. (2)* *A History of Psychology in Autobiography,* II, 1932, XIV. (3) *Occup. Psychol.* XXII, 1948, 20.
DREVER, JAMES (Junior) (1910——). (1) *Advancement of Science,* 13, 1956, 3.
EDGELL, BEATRICE (1871-1948). (1) *Brit. J. Psychol.,* 39, 1949, 121.
EDGEWORTH, FRANCIS YSIDRO (1845-1926). (1) J. M. Keynes, *Essays in Biography,* 1933, 267.
ELLIS, (HENRY) HAVELOCK (1859-1939). (1) *Human Affairs,* ed. by R. B. Cattell, John Cohen and R. W. Travers, 1937, 181. (2) H. Ellis, *Views and Reviews, 1st ser.,* 1932, frontispiece. (3) H. Ellis, *My Life,* 1940, frontispiece. (4) *Ibid.,* 148. (5) *Ibid.,* 288. (6) A. Calder-Marshall, *Havelock Ellis,* 1959, frontispiece. (7) *Ibid.,* 106. (8) J. S. Collis, *An Artist of Life,* 1959, frontispiece. (9) *Ibid.,* 56. (10) *Ibid.,* 72. (11) *Ibid.,* 168. (12) *Ibid.,* 200.
FISHER, RONALD AYLMER (1890——). (1) R. A. Fisher, *Contributions to Mathematical Statistics,* 1950, frontispiece. (2) *Sankhya,* 4, 1938, 264.
FLUGEL, JOHN CARL (1884-1955). (1) *Brit. J. Psychol.,* 47, 1956, 1. (2)* C. Murchison (ed.) *Psychologies of 1930,* 1930, XV. (3) *Advancement of Science,* 7, 1950-51, 103. (4) David Katz, *Psychologischer Atlas,* 1945, 91. (5) *Internat. J. Psycho-Anal.,* 37, 1956, 17.
FRAZER, JAMES GEORGE (1854-1941). (1) *Man,* 42, 1942, 1. (2) *Obituary Notices of Fellows of the Royal Society,* 3, 1939-41, 897.
FREUD, ANNA (1895——). (1) E. Jones, *Sigmund Freud: Life and Work,* 3, 1957, 144. (2) *Contemp. Psychol.,* 3, 1958, 83.
GALTON, FRANCIS (1822-1911). (1) J. P. Guilford, *Psychometric Methods,* 1936, VII. (2) David Katz, *Psychologischer Atlas,* 1945, 91. (3)* *Scient. Mo.,* 30, 1930, 124. (4)* H. E. Garrett, *Great Experiments in Psychology,* 170. (5) *Man,* XI, 1911, 33. (6) H. M. Walker, *Studies in the History of Statistical Method,* 1929, 175. (7) *Biometrika,* 2, 1902, frontispiece. (8)* K. Pearson, *Life, Letters, and Labours of Francis Galton,* III B, 1930, 663, index to 44 portraits.
GLOVER, EDWARD. (————) (1) O. Székely-Kovaks & R. Bérény, *Caricatures of 88 Pioneers in Psychoanalysis,* 1954.
GOSSET, WILLIAM SEALY ["Student"] (1876-1937). (1) *Biometrika,* 30, 1938, 205. (2) *Annals of Eugenics,* 9, 1939, frontispiece.
GORING, CHARLES BUCKMAN (1870-1919). (1) *Biometrika,* 12, 1919, 297.
GREENWOOD, MAJOR (1880-1949). (1) *Obituary Notices of Fellows of the Royal Society,* 7, 1950, 139. (2) *Biometrika,* 38, (Pt. 4), 1951, frontispiece. (3) *Brit. Med. J.* 1949, (15 Oct.), 877. (4) *Lancet,* 1949 (15 Oct.), 721.
HADDON, ALFRED CORT (1855-1940). (1) *Man,* 40, 1940, 96. (2) *Ibid.,* 42, 1941, 90. (3) A. H. Quiggin, *Haddon, the Head Hunter,* 1942, frontispiece. (4) *Ibid.,* 4. (5) *Ibid.,* 96. (6) *Ibid.,* 122. (7) *Ibid.,* 126. (8) *Ibid.,* 144. (9) *Obituary Notices of Fellows of the Royal Society.* 3, 1939-41, 449.
HALL, MARSHALL (1790-1857).* F. Fearing, *Reflex Action,* 1930, facing 123.
HARDING, DENYS WYATT (1906——). (1) *Advancement of Science,* 10, 1953-54, 101.

HARTLEY, DAVID (1705-1757). David Hartley, *Observations on Man, His Frame, His Duty and His Expectations*, 5th ed., 1810, frontispiece.
HAZLITT, VICTORIA (1887-1932). (1) *Brit. J. Psychol.*, 23, 1933, 205.
HEAD, HENRY (1861-1940). (1) *Obituary Notices of Fellows of the Royal Society*, 3, 1940-41, 665. (2) M. Critchley, *The Parietal Lobes*, 1953, 94. (3) *Brit. Med. J.*, 1940, (19 Oct.), 539. (4) *Lancet*, 1940, (26 Oct.), 534.
HEARNSHAW, LESLIE SPENCER (1907——). (1) *Advancement of Science*, 11, 1954, 119. (2) *Bull. Brit. Psychol. Soc.*, No. 29, (May 1956), centerpiece.
HICKS, G. DAWES (1862-1941). (1) *Proc. Brit. Acad.*, 27, 1941, 405. (2) A. Wood, *Bertrand Russell, The Passionate Sceptic*, 1957, 128 (group).
HOBBES, THOMAS (1588-1679). (1) F. Tonnies, *Thomas Hobbes, Leben und Lehre*, 1925, 3rd ed., frontispiece. (2) W. Molesworth, (ed.) *The English Works of Thomas Hobbes*, 1, 1829, frontispiece.
HOBHOUSE, LEONARD TRELAWNEY (1864-1929). (1) J. A. Hobson and M. Ginsberg, *L. T. Hobhouse; His Life and Work*, 1931, frontispiece. (2)* *Quart. Rev. Biol.*, 3, 1928, facing 502.
HUME, DAVID (1711-1776). (1) David Hume, *The Philosophical Works of David Hume*, 1, 1854, frontispiece. (2) J. Y. J. Greig, *The Letters of David Hume*, 1932, frontispiece and medallion. (3) J. Y. J. Greig, *David Hume*, 1934, frontispiece. (4) *Ibid.*, 281. (5) R. Klebansky and E. Moisner (eds.) *New Letters of David Hume*, 1954, frontispiece.
ISAACS, SUSAN (1885-1948). (1) K. S. Cunningham and W. C. Radford, *Education for Complete Living*, 1938, 337.
JACKSON, JOHN HUGHLINGS (1835-1911). (1) W. Haymaker (ed.) *Founders of Neurology*, 1953, 309. (2) K. Kolle, *Grosse Nervenarzte*, 1956, 136. (3) R. Dumesnil et al. *Les Médecins Célèbres*, 1947, 244
JONES, (ALFRED) ERNEST (1879-1958). (1) E. Jones, *Sigmund Freud, Life and Work*, 2, 1955, 64. (2) *Internat. J. Psychoanal.*, 37, 1956, 16. (3) W. B. Pillsbury, *History of Psychology*, 1929, 235. (4) *Brit. J. Psychol.*, 49, 1958, 177. (5) *Brit. J. Med. Psychol.*, 31, 1958, 71. (6) *Manchester Guardian*, 1958, (12 Feb.). (7) *Internat. J. Psychoanal.*, 39, 1958, 297. (8) *Ibid.*, 311. (9) E. Jones, *Free Associations*, 1959, 32. (10) *Ibid.*, 129. (11)* *Int. Zsch. f. Psychoanal.*, 15, 1929.
KLEIN, MELANIE (1882——). (1) M. Klein (ed.) *New Directions in Psycho-Analysis*, 1952, frontispiece. (2) *Internat. J. Psychoanal.* 33, 1952, 82. (3) N. Walker, *A Short History of Psychotherapy in Theory and Practice*, 1957, 104. (4) O. Székely-Kovaks & R. Berény, *Caricatures of 88 Pioneers in Psychoanalysis*, 1954, Caricature. (5) *Contemp. Psychol.* 3, 1958, 82.
KNIGHT, (ARTHUR) REX (1903——). (1) *Bull. Brit. Psychol. Soc.*, No. 20, 1953, centerpiece. (2) *Advancement of Science*, 5, 1948-49, 153.
LEWES, GEORGE HENRY (1817-1878). (1) A. T. Kitchel, *George Lewes and George Eliot*, 1933, frontispiece. (2) *Ibid.*, 204. (3) H. D. Traill and J. S. Mann, *Social England*, 6, 1904, 682. (4)* J. Sully, *My Life and Friends*, 1918, facing 262.
LOCKE, JOHN (1632-1704). (1) J. M. Baldwin, *History of Psychology*, 2, 1913, frontispiece. (2) J. Locke, *An Essay Concerning Human Understanding*, ed. by A. C. Fraser, 1, 1894, frontispiece. (3) J. Lough (ed.), *Locke's Travels in France, 1675-79*, 1953, frontispiece. (4) M. Cranston, *John Locke, A Biography*, 1957, frontispiece. (5) *Ibid.*, 97. (6) *Ibid.*, 113. (7) *Ibid.*, 224. (8) *Ibid.*, 240.
McDOUGALL, WILLIAM (1871-1938). (1) *Brit. J. Psychol.*, 29, 1939, 321. (2)* C. Murchison, (ed.) *Psychologies of 1925*, 1928, facing 273. (3)* C. Murchison (ed.) *Psychologies of 1930*, 1930, 17. (4) *A History of Psychology in Autobiography*, 1, 1930, 15. (5) R. S. Woodworth, *Contemporary Schools of Psychology*, 1948, 234. (6) A. A. Roback, *History of American Psychology*, 1952, Plate 19, 304. (7) *Human Affairs*, ed. by R. B. Cattell, John Cohen, R. W. Travers, 1937, 319. (8) *Obituary Notices of Fellows of the Royal Society*, 3, 39. (9)* *Ped. Sem.* 33, 1926, frontispiece. (10)* *Wittenberg Symposium*, 1928, 8. (11) David Katz, *Psychologischer Atlas*, 1945, 93. (12) H. L. Robinson, *William McDougall, A Bibliography*, 1943, frontispiece (from a bronze plaque). (13) *Man*, 42, 1941, 90. (14) R. B. Cattell, *Your Mind and Mine*, 1934, 104.

MACE, CECIL ALEC (1894——). (1) *Bull. Brit. Psychol. Soc.*, 1953 (May), centerpiece. (2) *Advancement of Science,* 8, 1951-52, 101.

MACDONNELL, W. R. (1852-1916). (1) *Biometrika,* XI, 1917, 281.

MALINOWSKI, BRONISLAW KASPER (1884-1942). (1) *Man,* 43, 1943, 1. (2) R. Firth (ed.) *An Evaluation of the Work of Bronislaw Malinowski,* 1957, frontispiece. (3) *Ibid.,* 38-39. (4) *Human Affairs,* ed. by R. B. Cattell, John Cohen and R. W. Travers, 1937, 199.

MARETT, ROBERT RANDULPH (1886-1943). (1) *Man,* 44, 1944, 33. (2) R. R. Marett, *A Jerseyman at Oxford,* 1941, frontispiece. (3) *Ibid.,* 310.

MAUDSLEY, HENRY (1835-1918). (1)* *J. Ment. Sci.,* 64, 1918, facing 117. (2)* *Mental Hygiene,* 14, 1930, suppl. following 462.

MILES, GEORGE HERBERT (1880——). (1) *Occup. Psychol.* 23, 1949, 193.

MILL, JAMES (1773-1836). (1) H. S. R. Elliott (ed.), *Letters of John Stuart Mill,* 1910, 103. (2) A. Bain, *James Mill,* a *Biography,* 1882, frontispiece.

MILL, JOHN STUART (1806-1873). (1) Michael St. John Packe, *The Life of John Stuart Mill,* 1954, frontispiece. (2) *Ibid.,* 128. (3) *Ibid.,* 504. (4) R. P. Anschutz, *The Philosophy of J. S. Mill,* 1953, frontispiece. (5) H. S. R. Elliott (ed.) *Letters of John Stuart Mill,* 1 & 2, 1910, frontispieces. (6) *Ibid.,* 2, 233. (7) H. D. Traill and J. S. Mann, *Social England,* 6, 1904, 419.

MILLER, EMMANUEL (1893——). *Human Affairs,* ed. by R. B. Cattell, John Cohen and R. W. Travers, 1937, 253.

MORGAN, CONWY LLOYD (1852-1936). (1)* *A History of Psychology in Autobiography,* II, 1932, 16. (2) *Obituary Notices of Fellows of the Royal Society,* 2, 25. (3) *Brit. J. Psychol.* 27, 1936, 1. (4)* *Quart. Rev. Biol.* 3, 1928, facing 448.

MYERS, CHARLES SAMUEL (1873-1946). (1) *Brit. J. Psychol.* 38, 1947, 1. (2) *Obituary Notices of Fellows of the Royal Society,* 5, 1945-48, 767. (3) *History of Psychology of Autobiography,* 3, 1936, vi. (4) *Occup. Psychol.* 21, 1947, 1. (5) *Man,* 42, 1941, 90. (6) *Lancet,* 1946 (26 Oct.), 622.

NUNN, T. PERCY (1870-1944). (1) *Brit. J. Educ. Psychol.* 15, 1945, 1.

PEAR, TOM HATHERLEY (1886——). (1) *Occup. Psychol.* 22, 1948, 107. (2) David Katz, *Psychologischer Atlas,* 1945, 94.

PEARSON, KARL (1857-1936). (1) *Obituary Notices of Fellows of the Royal Society,* 2, 73. (2) *Biometrika,* 28, 1936, 193. (3) *Ibid.,* 202. (4) *Ibid.,* 223. (4) *Biometrika,* 29, 1938, 161. (5) *Ibid.,* 183. (6) *Ibid.,* 235. (7) H. M. Walker, *Studies in the History of Statistical Method.* 1929, frontispiece. (8) *Ibid.,* 175. (9) K. Pearson, *Life, Letters and Labours of Francis Galton,* IIIA, 352. (10) *New Biology,* 1958, (No. 25), 9 (Caricature).

PHILPOTT, STANLEY J. F. (1888-1952). (1) *Brit. J. Psychol.,* 45, 1954, 244.

READ, CARVETH (1848-1931). (1) *Brit. J. Psychol.,* 22, 1932, 1.

RICHARDSON, LEWIS FRY (1881-1953). (1) *Obituary Notices of Fellows of the Royal Society,* 9, 1954, 217.

RICKMAN, JOHN (1891-1951). (1) O. Székely-Kovaks and R. Bérény, *Caricatures of 88 Pioneers in Psychoanalysis,* 1954, caricature.

RIVERS, WILLIAM HALSE RIVERS (1864-1922). (1) *Brit. J. Psychol.,* 13, 1932, 107. (2) W. H. R. Rivers, *Psychology and Politics and Other Essays,* 1923, frontispiece. (3) *Man,* 22, 1922, 97. (4) *Ibid.,* 41, 1941, 90. (5) A. H. Quiggin, *Haddon, the Head Hunter,* 1942, 96. (6) *Proc. Roy. Soc., Series B,* 95, 1923-24, 43. (7) A. C. Haddon, *History of Anthropology,* 2nd ed., 1945, 128.

ROBERTSON, GEORGE CROOM (1842-1892). (1) J. Sully, *My Life and Friends,* 1918, facing 182. (2) A. Bain and T. Whittaker, *The Philosophical Remains of George Croom Robertson,* 1894, frontispiece. (3) R. B. Perry, *The Thought and Character of William James,* 1, 1936, facing 596.

RODGER, ALEC (1907——). (1) *Bull. Brit. Psychol. Soc.,* 1, 1958, frontispiece.

ROMANES, GEORGE JOHN (1848-1894). (1) G. J. Romanes, *The Life and Letters of G. J. Romanes,* 1896, frontispiece. (2)* *Quart. Rev. Biol.,* 3, 1928, facing 488.

SCHONELL, FRED JOYCE (1900——). *Times Educational Supplement,* 1958 (21 March), 451.

SELIGMAN, CHARLES GABRIEL (1873-1940). (1) *Obituary Notices of Fellows of*

the Royal Society, 3, 1939-41, 627. (2) E. E. Evans-Pritchard (ed.) *Essays Presented to C. G. Seligman*, 1934, frontispiece. (3) A. H. Quiggin, *Haddon, the Head Hunter*, 1942, 96. (4) *Man*, 41, 1941, 133. (5) *Ibid.*, 42, 1941, 90.

SHAND, ALEXANDER FAULKNER (1858-1936). (1) *Brit. J. Psychol.*, 26, 1935, 323. (2) E. Westermarck, *Memories of My Life*, 1929, 73.

SHERRINGTON, CHARLES SCOTT (1857-1952). (1) *Brit. J. Psychol.* 44, 1953, 1 (2) *Obituary Notices of Fellows of the Royal Society*, 8, 1952-53, facing 241. (3) K. Kolle (ed.) *Grosse Nervenärzte*, 1956, 248. (4) Lord Cohen of Birkenhead, *Sherrington, Physiologist, Philosopher and Poet*, 1958, frontispiece. (5) *Ibid.*, 18. (6) *Ibid.*, 50. (7) *Brain*, 80, 1957, 402. (8) *Lancet*, 1952 (15 March), 569. (9) *Brit. Med. J.* 1952 (15 March), 606. (10) *Ibid.*, 1947 (22 November), 807.

SIDGWICK, HENRY (1838-1900). (1)* J. SULLY, *My Life and Friends*, 1918, facing 278. (2) A. S. and E. M. S[idgwick], *Henry Sidgwick, A Memoir*, 1906, frontispiece. (3) *Ibid.*, 16. (4) *Ibid.*, 303. (5) R. B. Perry, *The Thought and Character of William James*, 1, 1936, facing 596.

SMITH, GRAFTON ELLIOT (1871-1937). (1) *Obituary Notices of Fellows of the Royal Society*, 2, 1936-38, 323. (2) W. H. R. Rivers, *Psychology and Politics and other Essays*, 1923, frontispiece. (3) *Man*, 37, 1937, 41. (4) W. R. Dawson (ed.) *Sir Grafton Elliot Smith: A Biographical Record by his Colleagues*, 1938, frontispiece. (5) *Ibid.*, 151. (6) *Brit. med. J.*, 1937 (Jan. 9), 99. (7) *Lancet*, 1937 (Jan. 9), 113. (8) A. C. Haddon, *History of Anthropology*, 2nd ed., 1945, 120.

SPEARMAN, CHARLES EDWARD (1863-1945). (1) *Brit. J. Psychol.*, 37, 1946, 1. (1)* C. Murchison (ed.) *Psychologies of 1930*, 1930, xviii. (3)* *A History of Psychology in Autobiography*, I, 1930, xvi. (4) *Obituary Notices of Fellows of the Royal Society*, 5, 1945-48, 373. (5) David Katz, *Psychologischer Atlas*, 1945, 94. (6) J. P. Guilford, *Psychometric Methods*, 1936, viii. (7) *Brit. J. Educ. Psychol.*, 16, 1946, 1. (8)* *Wittenberg Symposium*, 1928, 9.

SPENCER, HERBERT (1820-1903). (1) J. M. Baldwin, *History of Psychology*, Vol. 2, 1913, 83. (2) D. Duncan, *Life and Letters of Herbert Spencer*, 1908, 26. (3) *Ibid.*, 77. (4) *Ibid.*, 127. (5) *Ibid.*, 392. (6)* J. Sully, *My Life and Friends*, 1918, facing 288.

STEWART, DUGALD (1753-1828). (1) Sir W. Hamilton (ed.), *The Collected Works of Dugald Stewart*, 1, 1854, frontispiece.

STOUT, GEORGE FREDERICK (1860-1944). (1) G. F. Stout, *God and Nature*, 1952, frontispiece. (2) *Proceedings of the British Academy*, 31, 1945, 307. (3) *Brit. J. Educ. Psychol.*, 16, 1946, 52.

STRAUSS, ERIC BENJAMIN (1894——). (1) *Bull. Brit. Psychol. Soc.*, 1956 (No. 29), centerpiece.

"STUDENT," see W. S. Gosset.

SULLY, JAMES (1842-1923). (1)* J. Sully, *My Life and Friends*, 1918, frontispiece. (2) E. Westermarck, *Memories of My Life*, 1929, 73.

THOMSON, GODFREY HILTON (1881-1955). (1) *Brit. J. Psychol.*, 46, 1955, 245. (2) *Brit. J. Educ. Psychol.*, 25, 1955, 65. (3) *Brit. J. Stat. Psychol.*, 8, 1955, frontispiece. (4) *Advancement of Science*, 1949-50, 170. (5) *History of Psychology in Autobiography*, 4, 1952, xii. (6) *Psychometrika*, 20, 1955, frontispiece.

TITCHENER, EDWARD BRADFORD (1867-1927). (1)* *Amer. J. Psychol.*, 38, 1927, facing 489. (2) *Ibid.*, 50, 1937, frontispiece. (3) *Ibid.*, 71, 1958, 10. (4)* *Studies in Psychology. Contributed by Colleagues and Former Students of Edward Bradford Titchener*, 1917, frontispiece. (5)* *J. Gen. Psychol.*, 1, 1928, frontispiece. (6)* *Scient. Mo.*, 26, 1928, 282. (7)* *Ibid.*, 30, 1930, 121. (8) R. S. Woodworth, *Contemporary Schools of Psychology*, Rev. ed., 1948, 21. (9) W. B. Pillsbury, *History of Psychology 1929*, 235. (10) *J. McKeen Cattell, Man of Science*, 2, 1947, 465. (11)* *Wittenberg Symposium*, 1928, 12.

TROTTER, WILFRED BATTEN LEWIS (1872-1939). (1) *Obituary Notices of Fellows of the Royal Society*, 3, 1939-41, 325. (2) E. Jones, *Free Associations*, 1959, 128.

TYLOR, EDWARD BURNETT (1832-1917). (1) *Man*. 17, 1917, 25. (2) R. R. Marett, *Tylor*, 1936, frontispiece. (3) *Anthropological Essays Presented to Edward Bur-*

nett Tylor in Honor of His 75th Birthday, 1907, frontispiece. (4) H. D. Traill and J. S. Mann, *Social England*, 6, 1904, 681. (5) A. C. Haddon, *History of Anthropology*, 2nd ed., 1945, frontispiece.
VALENTINE, CHARLES WILFRED (1879——). (1) *Brit. J. Educ. Psychol.*, 26, 1956, 1.
VERNON, HORACE MIDDLETON (1870-1951). (1) *Occup. Psychol.*, 22, 1948, 73. (2) *Lancet*, 1951 (24 Feb.), 477.
VERNON, MAGDALEN D. (————) (1) *Bull. Brit. Psychol. Soc.*, 1958, No. 36, frontispiece. (2) *Advancement of Science*, 16, 62, 1959, frontispiece.
VERNON, PHILIP EWART (1905——). (1) *Bull. Brit. Psychol. Soc.*, 1955 (No. 26), frontispiece. (2) *Advancement of Science*, 9, 1952-53, 91. (3) *Rev. Psychol. Appl.*, 4, 1954, 339.
WARD, JAMES (1843-1925). (1) *Brit. J. Psychol.*, 16, 1925, 1. (2) James Ward, *Essays in Philosophy*, with a memoir of the author by O. W. Campbell, 1927, 54, 59. (3)* *Amer. J. Psychol.*, 36, 1925, facing 449. (4)* R. H. Gault and D. T. Howard, *Outline of General Psychology*, 1925, facing 300. (5) R. B. Perry, *The Thought and Character of William James*, 2, 1936, facing 640.
WELDON, WALTER FRANK RAPHAEL (1860-1906). (1) *Biometrika*, 29, 1938, 193.
WOLTERS, ALBERT WILLIAM (1883——). (1) *Occup. Psychol.*, 22, 1948, 163.
WYATT, STANLEY (1890——). (1) *Occup. Psychol.*, 24, 1950, 65.
YOUNG, THOMAS (1773-1829). (1) A. Wood, *Thomas Young, Natural Philosopher, 1773-1829*, 1954, frontispiece. (2) G. Peacock, *Life of Thomas Young*, 1855, frontispiece. (3) *The Record of the Royal Society*, 3rd ed., 1912, 176, Plate XII.
YULE, GEORGE UDNY (1871-1952). (1) *Obituary Notices of Fellows of the Royal Society*, 8, 1952-53, 309.
ZANGWILL, OLIVER LOUIS (1913——). *Contemp. Psychol.*, 4, 1959, 99.

University of Manchester — JOHN C. KENNA
England

ON THE FLATTENING-EFFECT OF OPTICAL MAGNIFICATION

If one observes his surroundings through magnifying binoculars, the objects often appear to be distorted. Not only do they appear larger than they do when viewed with the naked eye, but they also appear nearer and somewhat 'flattened,' that is, foreshortened in depth. Bartley[1] and Miller and Bartley[2] have offered an account of the flattening-effect of binoculars by referring to the relation between the way in which the retinal images of the objects are changed by reduction in distance and the way in which the retinal images are changed by the optical system in binoculars. Essentially, the account involves a comparison of (1) the relation between the retinal images of the near and far edges of a cube, viewed at full distance vs. half-distance, with (2) the relation between the retinal images when the cube is viewed at full distance with the naked eye vs. a pair of 2× magnifying binoculars. Their account is summarized and interpreted in Fig. 1.

---

[1] S. H. Bartley, A study of the flattening effect produced by optical magnification, *Amer. J. Optom.*, 28, 1951, 290-299.

[2] J. W. Miller and S. H. Bartley, A study of object shape as influenced by instrumental magnification, *J. gen. Psychol.*, 50, 1954, 141-146.

A careful examination of this account, however, reveals that it is not in complete accord with certain aspects of optical theory. First of all, an assumption is made in the account that magnifying binoculars increase the retinal images of objects by an amount equal to the defined power of the system, regardless of the distance of the object, while reduction in object-distance alone increases the retinal images by an amount which depends on

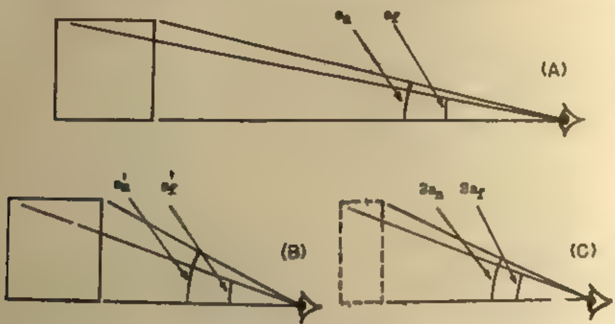


FIG. 1. SUMMARY AND INTERPRETATION OF BARTLEY'S ACCOUNT OF THE FLATTENING-EFFECT

The visual angles of the near and far edges of a cube are shown (A) when viewed with the naked eye at full distance and (B) at half-distance, as compared with (C) the image seen through a two-power telescope. If the distance of the cube is measured from the nodal point of the eye to the near edge of the cube, reducing the distance of the near edge by half doubles (approximately) the visual angle, so that $a'_n = 2a_n$. But since the far edge is reduced in distance by less than half, the angle a_f is not doubled, and $a'_f < 2a_f$. According to Bartley, a two-power telescope doubles the angles of both the near and far edges to yield $2a_n$ and $2a_f$, as in (C). The discrepancy between the disproportionate increase in the angles resulting from reduction in distance, as in (B) and the assumed proportionate increase in the angles by the optics, as in (C), is employed to account for the apparent flattening of the cube (see footnote 1).

the distance. While, however, the latter part of this assumption agrees with those of geometry, the former is not strictly in accord with optical theory.

A common definition of magnifying power employs the ratio of tangents of the visual angles subtended by object and image.[3] For small objects (less than about 3°) at great distances the ratio of tangents is approximately equal to the ratio of angles, and the definition in terms of the ratio of tangents may be interchanged with either the ratio of the angles themselves or the ratio of the focal lengths of the objective lens (f_o) and the eyepiece lens (f_e).[4] Thus, under these limiting conditions, the magnifying power (MP) may be defined as

$$MP = \tan\,\theta'/\tan\,\theta = \theta'/\theta = f_o/f_e \qquad [1]$$

Equation [1] may be used to find the size of the imaze of an object which is

[3] K. J. Habell and Arthur Cox. *Engineering Optics,* 1953, Ch. V. 185-205.
[4] *Ibid.,* 186.

'small' (as defined above) and viewed at a great distance.[5] If, however, the binoculars are focused on a nearby object of greater subtense than 3°, use of either the ratio of angles or the ratio of the focal lengths of the lenses would lead to an erroneous result. Under conditions other than the limiting ones above, an accurate statement of the size and distance of the image would require the use of the ratio of tangents,[6] or it would require a general formula which takes into account the changes which occur under non-limiting conditions. Southall provides a formula which does take these values into consideration:[7]

$$MP = \tan\Theta'/\tan\Theta = [f_e - Z(1 - cf_e)]/[U + f_o],$$

where $1/Z = z =$ distance of the image from the eye, $1/U = u =$ distance of the object from the objective lens. and $c =$ distance of the eye from the principal point of the eyepiece lens.

If the object is at a great distance and if the system is so adjusted that the image is at a distance equal to that of the object, Southall's equation reduces to a form equivalent to that of Equation [1]. If, however, object and image are not both at great distances, the actual distances would need to be determined to obtain a correct result.

Bartley's account of the flattening-effect also includes, implicitly, two assumptions about the location and size of the image formed by the optical system. The first is that the distance of the image does not change or is indeterminate; the second is that the size of the image (height) is constant.[8] These are made as parts of the original assumption that the proportional change in the retinal image formed by the optical system is the only variable of consequence in the flattening effect. The virtual image seen through the instrument is, however, of definite size and occupies a definite position in optical space.

The size and distance of the image formed by any optical system may be obtained from the general lens formula:[9]

$$1/D + 1/D' = 1/f \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [2]$$

where $D =$ object-distance, $D' =$ image-distance, $f =$ focal length of the lens.

By applying the lens formula successively to the objective lens and the eyepiece of a telescope the actual distance of the final image in optical space may be found. With some additional computations the actual size of the image may also be found.[10]

What has been said thus far suggests that an alternative account, in purely optical terms, may be offered for the flattening-effect. It is possible,

---

[5] R. L. Weber, M. W. White, and K. V. Manning, *College Physics*, 1952, 677-682.

[6] *Ibid.*, 677-682.

[7] J. P. C. Southall, *Mirrors, Lenses, and Prisms*, 1918, 460 ff.

[8] See Fig. 1. The flattened cube in (C) is obtained by projecting a line from the apex of the doubled angle back to a line parallel to the optical axis and at a height equal to the object height. A line is then drawn perpendicular to the axis to obtain the distance.

[9] A. C. Hardy and F. H. Perrin, *Principles of Optics*, 1932, Ch. IV. 63.

[10] The size and distance of the image will also depend upon the setting of the instrument, *i.e.* upon the positions of the lenses with respect to each other.

for example, that the distances and sizes of the images seen through a telescope are such that they are altered representations of the objects. If so, then it may be that a change occurs in perspective relations or in size-distance *gradients* rather than that a simple doubling of the visual angles occurs.[11] In the case of Bartley's example of the cube, it is possible that the *apparent* foreshortening may be directly correlated with an *actual* foreshortening of the image relative to the object. Such a correlation, if found, would provide a quantitative account of the flattening effect in precise psychophysical terms. A detailed examination of the physical properties of the

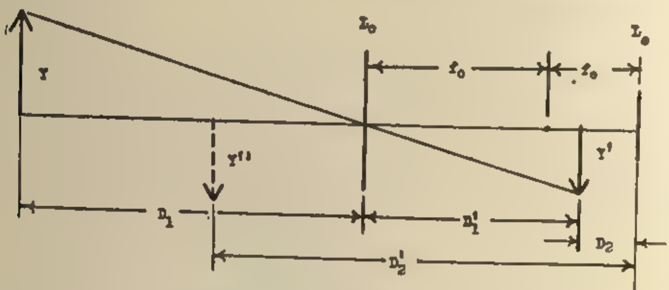


FIG. 2. GEOMETRY OF A SIMPLE TELESCOPE
(Not to scale)

The front lens (L_o) forms an image (Y') of the object (Y) at a distance (D'_1) behind the lens (L_o). This (real) image becomes the "object" for the eyepiece lens (L_e), which in turn forms a virtual image (Y'') at a distance (D'_2). The terms f_o and f_e refer to the focal lengths of the front lens and the eye-piece, respectively.

images seen through a telescope would provide a basis for such a quantitative account.

As an example of this kind of examination, let us observe what happens to the image of an object when it is viewed through a simple 2× magnifying telescope. For convenience, we have sketched the optical relations as shown in Fig. 2. The object, of height Y, is viewed at distance D_1. What will be the size (height) and distance of the image?

From the general lens formula above and from the properties of similar triangles, we obtain the distance (D'_1) and size (Y') of the image formed by the first lens. Since

$$Y/D_1 = Y'/D'_1 \qquad [3]$$

and

$$1/D_1 + 1/D'_1 = 1/f_o \qquad [4]$$

then

$$Y' = Yf_o/(D_1 - f_o) \qquad [5]$$

[11] J. J. Gibson, *Perception of the Visual World*, 1950, 116.

From Equation [4] we also obtain:

$$D'_1 = D_1 f_o/(D_1 - f_o) \ldots\ldots\ldots\ldots\ldots\ldots\ldots [6]$$

Equations [5] and [6] yield the values of size and distance of the image in terms of the size and distance of the object and the focal length of the objective lens. Similarly, we may find the size (Y'') and distance (D'_2) of the final image formed by the eyepiece. Replacing Y and D_1 in Equation [3] with Y'' and D'_2, applying Equation [4] and simplifying, we obtain:

$$Y'' = Y' f_e/(D_2 - f_e) \ldots\ldots\ldots\ldots\ldots\ldots\ldots [7]$$

and

$$D'_2 = D_2 f_e/(D_2 - f_e) \ldots\ldots\ldots\ldots\ldots\ldots\ldots [8]$$

To find the values of Y'' and D'_2 in terms of the size and distance of the object, we would need to know the setting of the instrument. If we assume that the focal points of the two lenses are coincident, that

$$D'_1 + D_2 = f_o + f_e, \ldots\ldots\ldots\ldots\ldots\ldots\ldots [9]$$

we may substitute the value of D_2 obtained from Equation [9] into Equation [8] to obtain

$$D'_2 = [(f_o + f_e - D'_1) \cdot f_e]/[(f_o + f_e - D'_1) - f_e] \ldots\ldots\ldots [10]$$

Substituting in Equation [10] the value of D'_1 obtained from Equation [6] and simplifying, we obtain:

$$D'_2 - [f^2_e/f^2_o]\,[D_1 - f_o] + f_e \ldots\ldots\ldots\ldots\ldots\ldots [11]$$

Since, from Equation [1], $f^2_e/f^2_o = 1/MP^2$, then

$$D'_2 = -[1/MP^2]\;\;[D_1 - f_o] + f_e \ldots\ldots\ldots\ldots\ldots\ldots [12]$$

the image being virtual and inverted.[12]

If a similar argument is applied to find the size of the final image, from Equations [3], [5], and [7], we obtain

$$Y'' = -\;Y/MP \ldots\ldots\ldots\ldots\ldots\ldots\ldots [13]$$

Equations [12] and [13] indicate that the final image formed by the telescope is not a magnified version of the object but is, in optical space, a minified version of the object. It is smaller by a factor of $1/MP$, and it is located at a distance nearer than that of the object by a factor of the order of $1/MP^2$. Thus, the image seen through a two-power telescope would be one which is half the size of the object and which is located at a distance of the order of one-fourth the distance of the object. While these two effects combine to produce a greater angular subtense and generally improved discrimination of detail, they may also combine to produce a considerably distorted appearance of size and distance.

---

[12] Hardy and Perrin, *op. cit.*, 479-480. In most binoculars the image is seen upright because some form of erecting system has been included in the design. If the image is erected by prisms, Equations [12] and [13] are adequate under the assumed conditions. If additional lenses are use to erect the image, their effects would need to be taken into account.

Optical space, in the case of a telescope, with properties assumed above, is a space which in fact is 'compressed,' not only in distance but also in size, relative to 'real' space. The relation between the variables of this optical space and the appearance of objects in it may be determined by psychophysical experiment. It is not unlikely that lawful relations will be found, as they have for 'real' space. The relation may perhaps be related to variables of gradient, as suggested above, or it may be that descriptions in terms of empirical laws of the kind suggested by Gilinsky, for example, will be made.[13] A detailed set of experiments, designed to produce the actual distortions of objects as they are produced by telescopes and arranged to determine the psychophysical relations between apparent size and distance and optical size and distance would be a necessary and important step in this account of the flattening-effect.

Washington University JAMES M. VANDERPLAS

A RUSSIAN REPORT ON THE POSTOPERATIVE NEWLY SEEING

Surgical acquirement of vision by the congenitally blind has always excited great interest. Not only are the accounts of how these people "learn to see" provocative, but the theoretical implications of the process discerned are of considerable significance.[1] Authentic instances of cases where the congenitally or near-congenitally blind have been rendered seeing are not many,[2] and those that have been reported have been the subject of dispute and interpretive disagreement.[3] For this reason, any addition to the literature reporting on restoration of vision to the congenitally blind supplements the meager stock of data extant and is, therefore, welcome.

In 1953 an article describing several cases of visual restoration appeared in *Vestnik Oftalmologii* [*Herald of Ophthalmology*], a Soviet journal.[4] While the account leaves much to be desired, it is unusual in some of its

[13] A. S. Gilinsky, Perceived size and distance in visual space, *Psychol. Rev.*, 58, 1951, 460-482.

[1] D. O. Hebb, *The Organization of Behavior*, 1949, 1-335.

[2] M. von Senden, *Raum- und Gestaltauffassung bei operierten Blindgeborenen vor und nach der Operation*, 1932, 1-303.

[3] M. Wertheimer, Hebb and Senden on the role of learning in perception, this JOURNAL, 64, 1951, 133-137.

[4] A. I. Pokrovskii, On the development of visual perceptions and judgments in the postoperative newly seeing in the light of the works of I. P. Pavlov, *Vestn. Oftal.*, 32 (6), 1953, 6-17.

detail and also in that the writer of the article himself, A. I. Pokrovskiĭ, Director of the Eye Clinic of the Voronezh Medical Institute, performed the surgery and conducted the postoperative studies and observations. In other words, the account is not second-hand, but stems from direct personal encounter with the cases reported.

Pokrovskiĭ first reports on the cases of two children, born with mild cataracts which, while sight-depriving, permitted reaction to light. Pokrovskiĭ records the age of one of the children as 10 yr. at the time of operation, but fails to mention the age of the other. He reports the following of their progress toward visual perception:

"1. After the operation both children were completely unable to recognize and name objects in their surroundings for some time, the duration of which varied for each. Their usual reaction was that they saw 'something,' but what it was, they did not know.

"2. Both quickly and easily recognized and named objects on running their fingers over them if these objects were of a kind frequently encountered in their home surroundings and, hence, familiar.

"3. If they looked [at] and simultaneously touched objects not ordinarily encountered in their home environment, but possessing forms and certain properties, tactilely discriminable and reminiscent of some familiar object or other, the children, on the basis of tactile contact, applied to the former names appropriate to the latter.

"4. For both children forms of objects were not visually discriminated, but were tactilely determined.

"5. Size and bulk of objects were similarly determined.

"6. The children were unable by vision alone to determine distance or, more exactly, the distance of the nearest objects. When walking, they collided with these objects and in the beginning employed the customary method of the blind to avoid them, putting their hands out in front and touching the objects in their way.

"7. Both children were completely unable by vision alone to determine the number of objects presented to them, but could do so by touch.

"8. Nor did they determine colors as such, distinguishing them only by brightness and naming one darker and another lighter.

"9. They possessed no conception of perspective, and their conceptions of extent and space, developed through touch and movement, were very limited in comparison with the same conceptions in normal children of the same age.

"10. In these newly seeing children, all the processes for discriminating objects, their form, size, etc., based in the beginning chiefly on the tactile-muscular analyzers [which include the receptors to the cortical termini], were without question carried out more slowly than in normal children operating by vision alone."

Pokrovskiĭ then considers the cases of two children who became blind in their early years and who had sight restored to them after an interval of five years. After stating that their postoperative progress to full vision was similar to that of the two cases reported on above, Pokrovskiĭ notes the following differences.

"1. Although the children did not recognize objects and reacted to them with 'I see something,' they did distinguish the size of this 'something' in comparison with an object of different size.

"2. They distinguished the colors of objects, though not too descriptively, and correctly named them.

"3. The most noteworthy feature of progress to fuller vision in comparison with that of the 2 born blind was its duration which was considerably shorter than in the case of the latter."

Pokrovskiĭ reports the following of the first case—a 9-yr. old girl, blinded at the age of five.

"She said that she saw everything, but in the beginning—and of this one can be easily convinced—she did not perceive what she saw; without question she did not recognize by vision alone the familiar objects of her surroundings, though she could quickly recognize and name these objects on running her fingers over them. In other words, she had to compare the visual sensations, which after long blindness were new to her, with the tactile-motor perceptions customary for the blind in order immediately to recognize and, one may say, 'see' the familiar object.

"When she was shown a kitten, she said, 'I see something gray; since I was a baby I have not seen anything like that.' On touching it, she exclaimed joyfully—"a kitten!" She did not recognize a dog, but said, 'I see something yellow.' On stroking the dog's back with her hand, she immediately named it. On seeing a cow, she said, 'I see something big, red and white.' On seeing a horse, she said, 'I see something big, gray and white.' She looked long at this and the other animal, listened to the particular sound of their breathing, and after some prying questions about animals' tails, faces, and the like, said "horse?" When she looked at a flower, she said, 'I see something green.' On touching the flower, she correctly named it as a flower.

"She did not recognize buckets of various sizes and color, but on handling them she correctly named them and correctly judged their differences in size and color.

"The girl could not judge the form of objects on the basis of merely sight, but on touching them did distinguish their forms. She distinguished the color and size of 2 little containers presented to her—one round, the other four-cornered—but she correctly determined their respective forms only by manipulating them.

"In the beginning the girl could not determine the distances to objects closest to her. All objects in front of her appeared very near to her. When requested to grasp any particular object, she extended her hand in the proper direction, but as a rule failed when trying to grasp even objects very near to her.

"The girl walked about her room during the first postoperative days very uncertainly, stumbling and knocking against objects standing in her way. But after several days she began to walk with steadily increasing confidence and after 2 weeks was already leading other patients around. Only when she found herself in unfamiliar surroundings, did she move around more circumspectly.

"She began rather quickly to recognize the objects about her, and at the time of discharge from the hospital, after 3 weeks, had begun without tactile intercession to utilize her vision, not in full, but in sufficient measure."

Pokrovskiĭ also refers to a second case—a girl of 8 yr., who was blinded

at the age of three. He reports that the progress toward visual perception was similar to that above, but very much slower. She was kept in the hospital for 2½ months and at the time of her discharge was not as advanced as the other girl although, as long as her surroundings were of the familiar kind, she was able to see and recognize things in her environment.

Pokrovskiĭ concludes by citing the cases of two patients who were blinded in their very early years and who were restored to sight much later in life—16 and 54 yr., respectively.

> "The 16 year old boy distinguished only light and grew up blind in a village as a member of a poor family living under very unfavorable conditions. Nobody had anything to do with him, and his mental development lagged greatly behind that of his seeing peers. His judgments of the surrounding world were elementary in the extreme. His behavior with respect to his surroundings before the operation reminded one of those born blind, and after a successful operation his development of visual perception also proceeded in accordance with that typical of the latter, but very slowly. While he recognized objects by fingering them, he was for a very long time unable to recognize them by sight alone. He did not have visual conceptions of size, form, bulk, color of objects, or their distance from him. He was unable to grasp an object in his visual field and erred many times in repeated attempts to accomplish this.
>
> "The patient was in the hospital for 2½ months after the operation and during this time learned to orient himself somewhat in his surroundings and to walk without stumbling against objects. He began to distinguish to some degree the form and size of objects and to distinguish their color not only by brightness but as more darkly or more brightly colored. However, even at the time of his discharge the visual perception of the patient was by no means as far advanced as in others with restored sight. From a letter, received from the patient 8 months after the operation, it was learned that he now goes about everywhere alone and distinguishes all colors of objects and that his vision has improved."

Concerning the patient who was blinded at the age of 2 or 3 yr. and who underwent surgery at the age of 54, Pokrovskiĭ then has this to report:

> "Before the operation he saw only light and did not distinguish color. He recognized his children and grandchildren by voice. Development of visual perception proceeded even more slowly than in the case just cited. Utilizing vision, he began to walk in the hospital with great difficulty and uncertainty. He was poor in recognizing people around him and, evidently, had frequent recourse to auditory cues in accordance with his previous wont. He would look for a long time at the face of his visiting son without recognizing him, but did so immediately at the sound of his voice. Unfortunately, it was not possible to keep track of this patient after discharge from the hospital."

Pokrovskiĭ discusses all of the above 6 cases in terms of Pavlovian theory and gives no indication of awareness of the differential importance of his observations and data for alternate theories. Pokrovskiĭ states that he and

others have personally dealt with a number of other such cases, but restricts his account to the above six as representative. Direct communication with the author on his report, though attempted, has not yet been established.

Brooklyn College IVAN D. LONDON

RED-GREEN COLOR DEFICIENCY AND COMPENSATORY LEARNING: AN EXPERIMENTAL CRITIQUE

In the March, 1957, issue of this JOURNAL, Overton and Brown report the results of an experiment which tested the implications of the hypothesis that "a permanent physiological weakness may elicit compensatory learning even though *S* has never been aware of the weakness."[1] They reasoned that because of this compensatory (and unconscious) learning, red-green color-weak *S*s who are not aware of their visual deficit will be more familiar with the spatial arrangement of the red and green lights on traffic-signals than *S*s with normal red-green vision.

Briefly, the investigators (1) determined that *S* was not aware of any color-vision deficiency; (2) tested the color-vision of 48 men and 24 women (presumably automobile drivers) using pseudo-isochromatic plates[2] that were illuminated by *incandescent* light[3] at an intensity level "slightly" below that recommended by the American Optical Company for use with the test; (3) asked each *S*, "On most traffic lights, is the red light or the green light on top?"; and (4) measured response-time (presumably with a stopwatch) between the last word ("top") of the question and *S*s' verbal reply, "red" or "green." The authors concluded that compensatory learning was demonstrated because significantly more 'color-weak' *S*s, *i.e.* those *S*s who missed one or more of the test plates, responded more quickly than normals (no plates missed) to the question about the traffic-light.

The purpose of the present paper is to criticize what appear to be several errors in this study and to summarize briefly the method and results of an experiment which was designed to eliminate these errors.

The presence of three important errors in Overton and Brown's experiment casts doubt upon the results and interpretation. First, by using a reduced intensity of incandescent illumination (instead of the recommended C.I.E. Source C) on the color-test plates, the investigators presumed that

[1] R. K. Overton and W. L. Brown, Unrecognized weakness and compensatory learning, this JOURNAL, 70, 1957, 126-127.

[2] American Optical Company, Pseudo-isochromatic plates for testing color perception, 1946.

[3] Personal communication with senior author.

'border-line' cases of color-deficiency would be discovered more readily. When, however, the test is used under these non-standardized conditions it is highly questionable whether red-green color-blindness is being tested.[4] The rather large proportion of *Ss* who made errors on test-plates lends support to this contention (27% men and 25% women missed two or more plates; 54% of the men and the women missed one or more plates). The generally accepted proportion of red-green deficiency in the population is approximately 8% for men and 0.4% for women.

A second error becomes apparent when one relates the method of analysis used by the investigators to the purpose of the color-vision test. The statistical analysis unquestionably indicates that the number of plates missed by *S* is directly related to the extent of his deficiency. This particular test, however, was, except for extremes of error-score, designed *solely* for screening and could not classify degree of deficiency.

A third criticism, perhaps not as crucial as the other two, refers to the possibility that the measurements of the dependent variable, *Ss*' response-times, were unwittingly influenced by bias on the part of *E*. The same *E* first administered the color-vision test and immediately thereafter measured the response-time to the question about the traffic-light.[5] Knowledge of *S*'s score on the visual test could influence *E*'s reaction-time and thus the measured response-time.

While the present experiment essentially repeats that of Overton and Brown, it differs in the logic of the proposition being tested and in the method of verifying this proposition. If it is true that very mild and unrecognized red-green color-blindness leads to compensatory learning of the orientation of lights on a traffic-signal, then red-green color-deficient *Ss who are cognizant of their deficiency* should also demonstrate greater familiarity with traffic-lights than do normal individuals. In other words, red-green color-blind *Ss* (with varying degrees of deficiency) who know they are color-blind should react faster than normals when asked about the spatial positions of the traffic lights.

To measure response-time accurately and objectively a standard method for determining choice reaction-time was used. The apparatus presented single stimulus-cards upon which were diagrammed a correct and an incorrect version of common highway-signals (sign of railroad-crossing and octagonal stop-sign used at intersections). *S* was instructed to press the

[4] L. H. Hardy, Standard illuminants in relation to color-testing procedures, *Arch. Opthal.*, 34, 1945, 278-282; L. H. Hardy, G. Rand, and M. C. Rittler, The effect of quality of illumination on the results of the Ishihara test, *J. opt. Soc. Amer.*, 36, 1946, 86-94.

[5] Personal communication.

button beneath the correct road-sign or signal as quickly as possible. The first two signs served primarily to permit *S* to become familiar with the task and apparatus and also served as a means of obtaining a measure of response-time which could be used as a reference-level in the analysis. The third stimulus-card, which will be referred to as Card 3, contained correct and incorrect diagrams of a stop-go traffic-signal. Instead of using red and green colors in the schematic diagram, the words STOP and GO were clearly printed across the appropriate 'light.' The *Ss*' color-vision was tested by means of pseudo-isochromatic plates in accordance with *all* recommendations stated in the instructions for this test (between 10 and 60 foot-candles of C.I.E. Source C).[6]

Forty men were used as *Ss* in the experiment. Twenty knew they were *not* color-blind and did not make an error on the four red-green screening plates of the visual test (normal group). The criteria for selecting the remaining 20 *Ss* were that each *S* knew he *was* red-green color-blind and in addition he had to make at least one error on the color-vision screening plates (color-blind group). All *Ss* were automobile drivers; the mean number of years of driving was 5.2 for the normals, and 5.9 for the color-blind.

As stated earlier, all 20 normal *Ss* responded correctly to the screening plates. In the color-blind group, the red-green defect of 8 *Ss* was categorized (according to the score-sheet of the American Optical Company) as mild, that of 9 *Ss* as medium, and that of 3 *Ss* as strong. Three *Ss* missed 5 or fewer plates (out of a possible 14 red-green test-plates), five *Ss* missed 7 to 9 plates and 12 *Ss* missed 10 to 14 plates.

Five color-blind *Ss* selected the incorrect diagram (STOP below GO) of Card 3 and their data were not used in the analysis.[7] Three of these *Ss* had a mild defect (having missed 5, 8, and 9 plates, respectively) and the remaining two had a medium defect (both missed 11 plates). Since the mean of the years of driving an automobile (6.8 yr., with a range from 3 to 15) was above the group mean, it is unlikely that this unfamiliarity with the traffic-signal was due to lack of experience in driving. This finding is not helpful to an interpretation in terms of "compensatory learning." The data of 7 normal *Ss* who had a mean of 3.6 yr. of driving and who made an incorrect response to Card 3 were also discarded.

Analyses of the data obtained from the remaining 28 *Ss* indicated (1) no evidence of a relationship between speed of response to Card 3

[6] American Optical Company, HRR pseudoisochromatic plates, 1955.

[7] Reaction-times to the incorrect stimulus on Card 3 were excluded from the statistical analysis because it is difficult, if not impossible, to evaluate the meaning of this type of response.

and degree of color-blindness (rho = .03, using only color-blind *Ss'* scores); (2) normal *Ss'* response-time to Card 3 did not differ statistically (median test, $\chi^2 = 0.57$, 1 *df.*) from color-blind *Ss'* when both groups' scores were corrected for initial differences in reaction-time by using the mean reaction-time for each *S* to Cards 1 plus 2 as a reference-value. Using only the uncorrected reaction-times does not change this finding (median test, $\chi^2 = 0$).

If the criticisms of the study of Overton and Brown are valid and the results of the present experiment are free from similar errors, then there is good reason to believe that very mild or strong red-green visual deficiency does not result in a "compensation" for this deficiency by the learning of the position of the signal-lights. This conclusion is not at all surprising when consideration is given to the well-known fact that usually a 'green' traffic signal is really bluish-green and the 'red' is really yellowish-red.[8] In fact, several of the red-green deficient *Ss* who served in the study were surprised to discover at the end of the experiment that the investigators ever doubted their ability to distinguish between the two signal-lights. These *Ss* stated that they did not have to learn the positions of the signals because it was so simple to 'see' that the red light was of a different color from the green.

Thus, the experiment of Overton and Brown, although ingenious in its conception, was faulty in its method. Their results do not unequivocally support the hypothesis that an unrecognized physiological weakness may elicit compensatory learning.

University of Missouri

ALVIN G. GOLDSTEIN
C. ROBERT BORRESEN

THE SENSORY RANGE OF ELECTRICAL STIMULATION OF THE SKIN

Application of alternating current to the skin under appropriate conditions will elicit sensations of pain, pressure, tingle, warmth, and cold. Of these, however, only pain and tingle can routinely be elicited. To determine the limits of the range of physical intensities useful for cutaneous stimulation, it is necessary to determine the amount of current required to reach absolute threshold (*RL*) of tingle and of pain and the tolerance-limit (*TL*) of current-intensity.

The study reported in this Note was undertaken to determine the values of current required for the *RLs* of tingle and of pain and for the

[8] American Optometric Association, *Manual on Drivers' Vision Test*, 1949, 25-26.

TL. Two highly trained *O*s were used. After preliminary training, each *O*'s responses were recorded during a period of 10 sessions with sine-wave AC at each of the following frequencies, 100, 500, 1000, 5000, 10,000 ~. After determining the *RL* of tingle, the current was increased until *O* reported the appearance of pain. This value was recorded and then the current was increased still further until *O* reported the *TL*. Occasional *Vexierversuche* were given but no false reports were made.

Mean values of current required for the *RL*s of tingle and of pain and the *TL* of current-intensity are plotted in Fig. 1. In all instances, more current was required for reports of pain than for tingle, and more

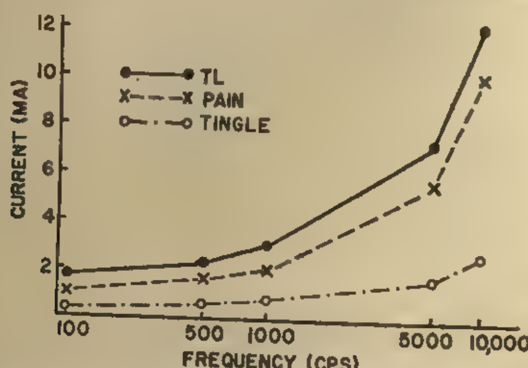


FIG. 1. THE *RL*s AT VARIOUS CURRENT-FREQUENCIES OF TINGLE AND OF PAIN AND THE *TL* OF CURRENT-INTENSITY
(Mean for 10 sessions and 2 *O*s.)

current was necessary for the *TL* than for the *RL* of pain. More current was, furthermore, necessary to reach the thresholds at the higher frequencies of stimulation than at the lower. The functions were consistent (see Fig. 1) for the various thresholds for both *O*s, *i.e.* a positively accelerated ascending curve.

Relative reliability of the reports of tingle, pain, and the *TL* was determined by Bartlett's test of homogeneity of variance. The results indicated significantly (1% level) different distributional variances except for one *O* at 5000 ~. Examination of the variance-distributions indicated that the *RL* of pain was the most variable for many, but not all, frequencies. To determine whether the variances were significantly different for the various frequencies, sensations, and *O*s, an analysis of variance was made. The results showed no significant effects, but significant interactions indicated that the magnitudes of the variance changed for the various sensations at different frequencies and with different *O*s.

Qualitative reports of the sensations elicited by electrical stimulation indicated that tingle was typically a weak sensation localized in a small area under the active electrode. Pain was typically a localized sensation superposed on the tingle, and described as similar to the sensation elicited when a needle point penetrated the skin. The *TL* typically was reported to be like sensations of burning and muscle contractions and of an apparent intensity which could not long be endured with equanimity. Little or no change in the quality of sensation was reported with change in frequency of stimulation, except that the apparent rate of vibration was obviously slower than the physical rate of change at 5000 and 10,000 ~.

In all but one instance, at a stimulation frequency of 10,000 ~, the ratios of *TL* to the *RL* reported by Anderson and Munson are considerably larger than the ratios of the *TL* to the *RL* of tingle found in the present study. The different ratios of the two studies could be due in part to the fact that the *RLs* of tingle were recorded in the present work, whereas Anderson and Munson report current-values required for the "threshold of sensation."[1] Their "threshold" may have been touch or tingle. Exploratory work has indicated that touch can often be elicited at lower current-values than those required for tingle. The site of stimulation in Anderson and Munson's experiment was the forearm, a site where higher current-values at the *RL* might reasonably be expected than at the more densely innervated area of the present study, the finger tip; comparison of the data from the two studies indicates the reverse.

The *RLs* of pain obtained in the present study tend to be somewhat lower than those of Lindner, due perhaps to different subjective criteria for the reporting of pain. Lindner's *Os* reported pain to have "stretching" or "burning" characteristics.[2]

The *Os* in the present experiment were able to report the tolerance-limit (*TL*) as reliably as the *RL* of pain and tingle, and after preliminary training evidenced no undue emotional reactions to stimulus-intensities producing pain. If an investigation is made using stimulus-intensities that elicit pain, prior training will probably be necessary to avoid emotional reactions. It should be noted that the energy ranges determined by such a procedure are still considerably less than comparable ranges for vision, audition, or mechanical vibration of the skin.

U. S. Army Medical Research Laboratory

G. R. Hawkes
J. S. Warm

[1] A. B. Anderson and W. A. Munson, Electrical excitation of nerves in the skin at audiofrequencies, *J. acoust. Soc. Amer.*, 23, 1951, 155-159.

[2] R. Lindner, Physiologische Grundlagen zum elektrischen Sprachetasten und ihre Anwendung auf den Taubstummenunterricht, *Z. Sinnesphysiol.*, 67, 1937, 114-144.

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Psychology: A Study of a Science. Study I. *Conceptual and Systematic.* Volume 1. *Sensory, Perceptual, and Physiological Formulations.* Edited by SIGMUND KOCH. New York, McGraw-Hill Book Company, Inc., 1959 Pp. x, 710. $9.75.

This is the first of seven volumes growing out of "Project A," a study initiated by the Policy and Planning Board of the American Psychological Association. The purpose of the project is to analyze the "methodological, theoretical, and empirical status of psychological science." The execution of the project has been mainly the responsibility of the Editor, although he has been advised by a panel of consultants and has received overall guidance from a steering committee. *Study I* comprises the first three volumes of the series and is devoted to a survey of fields and systematic formulations within psychology. *Study II,* in Volumes, 4-6, is to be devoted to "Empirical Substructure and Relations to Other Sciences." Finally, Volume 7 is to be Dr. Koch's own views growing out of the study and titled "Psychology and the Human Agent."

This review will be concerned mainly with a total impression of Volume 1. A detailed criticism of the nine individual contributions contained in this volume would require a number of specialists. Such piecemeal criticism does not seem desirable for two reasons: first, the individual contributions are written by men of undoubted authority in their fields, and secondly, the justification for the volume lies in its total impact, rather than in the individual excellence of the papers.

The evaluation of the whole project will have to await the completion of the series of publications. Nevertheless, the present volume cannot be treated independently of the over-all plan of the series, since it represents an important application of the plan, and since also it contains an introduction outlining the purposes of the series as a whole. Even after reading Volume 1, this reviewer is extremely puzzled about the purpose of undertaking the whole project. The original statement of the aim of "Project A" and the title of the series of volumes both seem to imply a hope of discovering organizing themes for psychology as a whole. On the other hand, Koch disclaims any such expectation. Instead he insists that psychology consists of the efforts of many individual men, working with different concepts and organizing frameworks. As he puts it: "Each actor in a science acts from and in his own image of that science. The happy paradox of science is that the vision, image, must be his *own,* yet partly formed and tested by the visions of others. A vision of a science must be unique, private: yet that view will not be *in* science unless fed by other viewers, present or past. The assemblage, then, of multiple expert views of common themes and their publication can enlarge, sharpen, or confirm the private vision of any individual reviewers of these views. The reader—student, investigator, analyst, or civil man—may confront these assembled views and see a total pattern, or a selective one of any conformation" (p. 8).

What, then, is the advantage of assembling in this special series of volumes a num-

ber of these individualized views of men working on many different problems in psychology? Couldn't the reader get the same effect by surveying recent issues of various psychological journals? Apparently, Koch and the committee felt that specially prepared contributions, in which each author was to emphasize his modes of thinking about his field and the influences which he felt to be important in his thinking, would represent the state of psychology better than the usual professional publications of the same authors.

To produce this emphasis upon modes of thinking, the authors were provided with a list of "Suggested discussion topics." Among the twelve subdivisions of the list are such headings as "Background factors and orienting attitudes," "Structure of the system as thus far developed," (including questions concerning independent and dependent variables, etc.), "Initial evidential grounds for assumption of system," "History of system to date in mediating research," etc. The authors were not constrained to follow this outline, however, and did not even have to refer to it. The results vary from those who follow the discussion topics point for point to those who refer to the outline only as an afterthought.

In spite of this procedure, this reader cannot discover in himself any new impressions of the nature of psychological thought or research, nor can be believe that younger and less experienced students would gain any more than they would from an equivalent selection of readings from the general literature in the journals. This reviewer was substantially enlightened about some fields in which he had not been keeping up with the current literature, and he found his understanding of issues sharpened in other areas where he was more conversant with the literature. He has profited substantially from the book in these ways, but these, presumably, are not the main aims of the project.

The above conclusions may appear to prejudge the whole series on the basis of a single volume, and may also appear to be unduly severe criticisms of the work of Dr. Koch and his collaborators. As for the first point, I feel that the fields of sensory, perceptual and physiological conceptions should provide a broad enough test of the overall value of the aims of the project, but I also recognize that these first impressions may have to be revised when the series is completed. As for the possibility that the criticism is too severe, I can only say that the fault was in the original concept of Project A and not in the execution. Koch and his authors were given what seems to me either a meaningless or an impossible task. In view of this situation, the results are surprisingly useful. Above all, Koch has not attempted to force the whole field into an artificial logical or methodological mold which would have made the result dangerous rather than merely vague in purpose.

The final conception of the task consisted, as I have indicated, in an extended series of exhibits of the systematic thinking of distinguished psychologists about their own fields of specialization. The result is that the volume is more likely to be sampled by the general psychological reader than to be read as a whole. While the *Study* is not intended to be a handbook in the sense of a systematic survey of facts, method, and principles, many of the chapters will probably become required reading for coming generations of graduate students who are seeking a full understanding of particular topics. In keeping with this possible use of the *Study*, I shall give a brief survey of the specific contribution is to be found in Volume 1.

"Three Auditory Theories" by J. C. R. Licklider deals with three separate problems —the detection of the presence or absence of a signal, the intelligibility of speech,

and the mechanism of pitch perception. Of the three, the last is Licklider's own development, while the others were mainly the work of other authors. Signal detection is dealt with in terms of statistical decision-theory. Intelligibility is related mathematically to the energy distribution of the speech signal and background noise. Licklider's "triplex theory" of pitch perception is a model of the physiological mechanisms in terms of their functional mathematical properties, but based upon recent knowledge of cochlear function. This is a chapter for the serious and advanced student of the field who has a sound mathematical and physical background.

"The Quantum Theory of Light and the Psychophysiology of Vision," by M. H. Pirenne and F. H. C. Marriott is concerned with the process by which light stimulates the retina. On the psychological side, the theory is used to predict the results of experimental studies of the absolute threshold for for brightness. The theory is derived, however, primarily from physical and photochemical considerations. Response curves are predicted from various hypotheses concerning the number of quanta necessary for stimulation of retinal units, and the curves compared with empirical results. Such studies also have implications for the independence or dependence of retinal units, and the separation of variability of response due to biological or internal factors from that due to statistical variations in the stimulation process itself.

"Theory of Stereoscopic Vision," by Kenneth N. Ogle describes present knowledge about the functional relationships between perceived depth and retinal disparity (under the abstracted conditions of the stereoscope). The main variables are those which were dealt with by Wheatstone, Helmholtz and Hering, and no radically new concepts seem to have been developed. Ogle has presented a useful summary of present knowledge of the experimental findings on this problem. It probably would have been clearer if he had not felt constrained to follow so closely the "suggested discussion topics" of the editor. The "theory" of stereoscopic vision is little more than a first-order relationship between stimulus-conditions and reported depth, whereas the discussion outline is aimed at a more elaborate "system" of concepts.

"The Luneburg Theory of Binocular Space Perception" by Albert A. Blank is a description of the mathematical properties of "visual space," which is found to be "locally euclidean" but more generally it is a hyperbolic riemannian space. The theory is explicitly limited to the binocular perception of points of light in the dark, and its importance for normal space perception with multiple cues is unknown.

"The Systematic Psychology of Wolfgang Köhler" by W. C. H. Prentice is the first of four chapters dealing with more general approaches to perception at large. It deals with general trends and issues rather than the detailed analysis of a concrete problem which is characteristic of the earlier chapters in the volume. It summarizes the general characteristics of gestalt psychology as well as Köhler's particular research and contributions. The article requires less specific technical competence on the part of the reader, but, since it is so general, its full value still depends upon a baskground of familiarity with more specific analyses and research in perception. It is a sympathetic exposition, but not a polemic for the gestalt point of view.

The next two papers present sharply contrasting points of view ("Perception as a Function of Stimulation" by James J. Gibson and "Brunswik's Probabilistic Functionalism" by Leo Postman and Edward C. Tolman). Both are concerned with the overall relationship between stimulation and the perceived result, and both deal with the perception of objects in a complex milieu as contrasted with the artificial situation of the stereoscope or points of light in the dark. Both have a point of view on the role

of learning in perception. Gibson, however, accounts for the properties of the perceived world in terms of psychophysical correspondence to "higher order" patterns in the stimuli. Thus, size-constancy occurs as a result of the orderly relationship between the size of the projection of an object and "the size of the elements of texture or structure in the adjacent optical array." (p. 479). For Gibson, learning is a process of discriminating relationships which are already "there" in the stimulus-array. To Brunswik, on the other hand, stimulus-cues are merely signs which are more or less probable indicators of the objective properties of the world. The emphasis is upon learning the probable validity of the cues, rather than their inherent orderly relation to to the perceived result. In this volume the views are set side by side, with little direct comparison or confrontation of their differences. The reader is therefore urged to supplement these accounts by reading Hochberg's brilliant comparison of the two views.[1]

"Adaptation-Level Theory" by Harry Helson presents an organizing principle which Helson regards as uniting many different psychological phenomena, from the perceived hue of objects in colored illumination to personality and social behavior. In its broadest application, the principle becomes a statement that the effects of any particular stimulus or stimulus-object are a function of other surrounding or background stimuli, and of the past history of stimulation for that particular organism. For some problems of perception, Helson has worked out quantitative statements which allow the computation of the adaptation-level as a composite effect of all of these background stimulus factors. In the more general application, however, actual computation of the adaptation-level is no longer possible, and the concept is used only as a general framework for thinking about the problems. Again, it would be very helpful if there were more cross-comparison relating this point of view to the others presented in the volume.

"A Neuropsychological Theory" by D. O. Hebb appropriately stands at the transition from studies of perception to physiological conceptions of psychology—appropriately in that perception plays a central role in the theory and also in that his conception of neural function is built upon perceptual and behavioral studies. The theory is presented in highly condensed form, because Hebb has elected to concentrate upon the editor's discussion-outline rather than to expound the theory within its own framework. Thus the student may consult this article for background of the development of Hebb's theory, but he will find other sources more useful for understanding the theory itself.

"Physiological Theory of Drive" by Clifford T. Morgan summarizes the status to date of his theory that drives consist of, or depend mainly upon, states or activities within the central nervous system. The structure of the theory is based mainly upon studies relating to the viscereal drives, notably hunger and thirst. Morgan discusses recent studies dealing with exploration, manipulation and curiosity, suggesting that the same general theory may be applied to these drives as to hunger and thirst, although physiological evidence is virtually non-existent. The general graduate student will find this a useful and clear summary of Morgan's position.

I have tried to summarize the separate papers with as little evaluative judgment as possible. The composite impression is of general excellence of exposition, but there

---

[1] J. E. Hochberg, Review of Egon Brunswik's *Perception and the Representative Design of Psychological Experiments*, this JOURNAL, 70, 1957, 480-485.

seems to be little relationship between the excellence of the paper and its use of the Editor's "Suggested Discussion Topics." In some cases the topics seem appropriate to the area under discussion and provide a useful framework for the author, in others the topics seem inadequate for the kind of theory involved, or fail to fit the author's ways of thinking and organization. In the latter cases some authors wisely go their own ways, others try to follow the outline and fail to do justice to their ideas.

In judging the overall effect, we must note that Koch has disclaimed any attempt to provide a representative sampling of systematic positions, and points out that some areas are inadequately represented because of the inability of appropriate authors to participate. He notes especially the scanty coverage of the physiological conceptions, and we can agree that this is an unfortunate deficiency. The picture of a developing science is also incomplete in that it fails to include some of the thinking of younger and less well-known men in these fields, but this is also a defect which would be difficult or impossible to remedy.

We can agree with Koch that different approaches and concepts are needed for different problem areas. We can also agree that it is desirable and characteristic for different psychologists to develop different approaches and frames of reference even within the same problem-area, and that progress will be fostered by this diversity. As Koch puts it, the picture of a science will necessarily be a "collage" rather than a "colligation." It is possible, however, to overemphasize the diversity and to miss the degree to which differences are merely matters of verbal preference, or differences in the aspects of the field which are dealt with, rather than truly contradictory in that they produce different predictions for the same conditions.

In this volume there are six papers dealing with the same or overlapping problem areas in perception. The theories or points of view in these papers have been presented before in the literature, perhaps not so well nor so systematically, but still the elements could be found in previous works. If more space had been used in analyzing systematically the interrelationships of these points of view and less in direct exposition, our understanding of the status of this problem area could have been substantially advanced. The 'discussion-outline' 'might have provided more materials for such a cross-comparison if all of the writers had used it in comparable ways, but the major work of sorting and comparing would still have been left to the reader. Only occasionally does one of the authors deal directly with the ideas in one of the other systems and give in detail his objections or the extent to which he agrees or disagrees on specific points.

In short, this volume is valuable for its excellent parts, but this is at least one instance in which the whole is not more than the sum of its parts. T.A.R.

Mathematical Puzzles & Diversions. By MARTIN GARDNER. New York, Simon and Schuster, 1959. Pp. xi, 179. $3.50

In *Mathematical Puzzles & Diversions,* Martin Gardner, long a contributor to the *Scientific American,* brings together, with some new materials, the best recreational 'brain-teasers' that have appeared over the years in that magazine.

This book, a pot pourri of puzzle lore, ranges, in its short chapters from paper folding (the flexing of hexaflexagons), a diversion developed in 1939 by English and American mathematicians, to the problem, in Chapter 16, of right-left in mirror reflections and of palindromic words and sentences. The intervening chapters are

devoted to magical matrices, probability paradoxes, curious topological models including Moebius strips and the Klein bottle, polyominoes, mathematical fallacies, and card tricks. One chapter is devoted to the best of Sam Loyd's puzzles, one to mnemonic methods of memorizing numbers, and others to describing and explaining the mathematical principles involved in the games of Ticktactoe, Mill, the Tower of Hanoi, Nim, Tac Tix, the ancient Japanese or Chinese game of Go-moku, and the game of Hex developed a few years ago in Niels Bohr's Institute for Theoretical Physics in Copenhagen. Mechanical and wire puzzles were not touched upon in this book, though they may properly, through the geometry of space perception, be regarded as mathematical.

The puzzle enthusiast will meet many old friends in the pages of this book. Most of them, however, are accompanied by interesting commentaries and adorned with new twists and turns that Gardner has collected from the literature and from his correspondents over the years.

For example, the old problem (p. 23) of the hunter who "walks a mile due south, turns and walks a mile due east, turns again and walks a mile due north where he finds himself at his starting point. He there shoots a bear. What color is it?" The time honored answer is "white" because the only point upon the earth that satisfies these spatial and biological requirements is the north pole. Insofar as the spatial conditions alone are concerned, there are many points upon the earth that satisfy them. Can you, Gardner asks his readers, think of one? There are an infinite number, but it is not until the questionee escapes the determining tendency established in the original problem by the habitat of the white bear that he is able to think of the antarctic and thus arrive at the solution.

Again, in Sam Loyd's superb threefold chess problem, known the world over as the problem of "Charles XII at Bender," Gardner adds a little-known variation that surpasses the other three. According to the story with which Loyd dressed these problems, King Charles of Sweden was playing chess with his Prime Minister while his army was being beseiged by the Turks at Bender. Near the end of the game Charles announced, "mate in *three* moves." At this utterance, a stray bullet shattered his knight. While his opponent sought to replace the knight, Charles studied the board and said that he did not need the piece because he could still mate in *four* moves. At this announcement, a second bullet removed one of his pawns. Still unperturbed, Charles announced that he could now mate in *five* moves. Loyd's problems were to discover these successive mates. To these problems and story, Gardner adds the variation published by Friedrich Ameling, a German chess expert, in the *Baltische Schachblätter* in 1900 (p. 274); namely, that, had Charles' rook instead of his knight been removed by the first bullet, he could still have mated his opponent in *six* moves —a variation that had escaped Loyd's notice and that of other chess experts for 40-odd years! Though reported in the chess literature, Ameling's variation is seldom reported in the puzzle literature.

Loyd's famous (or infamous, as you will) "14-15" or "Boss" puzzle, known also in Europe as "Le jeu de taquin," was composed in 1878. It consists of a shallow, square box which contains 16 square spaces. Into these, 15 little square blocks, numbered from 1 to 15, are placed, the lower right-hand corner being left vacant that the blocks may be moved about and their positions interchanged. At the beginning, the blocks are arranged from left to right in serial order with the exception of the last two, which are reversed, *i.e.* block 15 is placed before block 14. The problem is to slide the blocks about until all of them are arranged in serial order, that is, block 14 is placed in its proper position before 15. Apparently a simple problem! For several years after its appearance, the puzzle enjoyed great popularity throughout Europe as well as America. Loyd offered a prize of $1000 for the correct solution, and, as Gardner reports, "Thousands of people swore they had solved it, but no one could recall his moves well enough to record them and collect the prize."

Loyd was safe because his problem, as two American mathematicians (W. W. Johnson, Notes on the '15' puzzle, *Amer. J. Math.*, 2, 1878, 397-399; W. E. Storey, *idem*, 399-404) proved mathematically, was not solvable. The number of possible positions of the blocks within the box is 16!, *i.e.* nearly 21 trillion (20,922,789,888,000). Half are of one 'parity' and half of another. All positions in one parity may be reached from that parity but a position of one parity can not be reached from any position of the other parity. The arrangements of the blocks "14-15" and "15-14" are in different parities, hence one cannot be reached from the other.

Despite the exposure of this hoax, the "Boss" puzzle has a resurgence with every new generation of puzzle enthusiasts and it may still today be purchased at the toy counters of most notion stores. It is highly probable, however, that Loyd was not enough of a mathematician to know what he had perpetrated upon his public as he was too much of a puzzlist to break the cardinal principle of every puzzle; namely, the capability of its being solved fairly without any tricks or fraud. The exposé of a hoax in a scientific magazine is effectively buried as it seldom reaches the eyes of those concerned.

The diversions discussed in 11 of the 16 chapters of this book are explained in their description and exposition. In only 5 chapters are the problems left for the reader to solve. In these instances, the answers are not far to seek as they are given in a section immediately following the statement of the problems. Most of the chapters are followed by Addenda which give interesting side lights and commentaries upon the subjects treated in the preceding chapters that Gardner culled from the literature or received from his correspondents.

References for further reading upon the subjects of the various chapters and a very brief (11 lines) statement "about the author" close the little volume. The lack of an index of the topics treated and of the puzzles posed is regrettable. The interested reader, unless he prepares his own index, cannot find the problems that intrigued him without paging through the book.

The University of Texas — CHARLES VALLEY BROOK

Educational Psychology. By GEORGE G. THOMPSON, ERIC F. GARDNER, and FRANCIS J. DI VESTA. New York, Appleton-Century-Crofts, 1959. Pp. xx, 535. $6.00.

The authors explain in the Preface that their goal in the planning of this book was to produce "a truly functional interpretation of the most useful generalizations that modern psychology can offer the classroom teacher." The present reviewer believes that in most respects the book lives up to the authors' hopes for it.

The selection of content and the manner in which it is presented in this book will best be understood by the reader if he translates "classroom teacher" to mean prospective teacher. It is the undergraduate with little previous work in psychology for whom this book was written. This may be inferred from a statement by the authors in the Preface regarding their way of organizing the book's content. The statement is, "This approach presents the principles of developmental psychology from a new perspective to the student who may have taken a prior course in human growth and development and as [*sic*] a stepping-off point for all students."

Especially commendable is the authors' anticipation of the possibility that in some schools their *Educational Psychology* would best serve as the textbook in a course for which another (in human growth and development) was the antecedent. The 'approach' mentioned in the quotation is well adapted to such a sequence of courses. First, it provides for an overview of the influences of maturation and

environmental forces in the individual's growth and development—with the implications spelled out for the teacher. Secondly, it provides for the integration of these influences into the discussion of later chapters dealing respectively with evaluation, learning and adjustment.

The chapters of the book have been grouped under five main concerns expressed, respectively, in the titles of the main parts of the book: I. An Overview of Growth, Learning and Adjustment of Your Pupils; II. Understanding the Individual Pupils in Your Classroom; III. Achieving Optimal Learning Conditions in Your Classroom; IV. Guiding Your Pupils Toward Effective Adjustment in Everyday Living; and V. Your Mental Health as a Teacher. These parts and the chapters included in them correlate well with the selection and sequence of topics found in numerous introductory courses in educational psychology.

The book consists of 24 'relatively independent' chapters. Their independence and their limited length (they average about twenty-five pages) give the book considerable flexibility and adaptability. Instructors who use this textbook will probably find, as the authors hope, considerable freedom in relating their own course outline to it.

Accompanying the textbook are a *Student Workbook* and an *Instructor's Manual* which offer, respectively, a variety of exercises and projects for the student and suggestions of evaluative and audio-visual aids for the teacher.

An especially pleasing feature of this book is the clarity with which the main ideas are presented. In this respect the authors have succeeded notably in making the generalizations of modern psychology understandable to prospective teachers. With remarkable regularity, a clearly worded generalization is the opening sentence of key paragraphs. Many of these generalizations amount to statements of principle. Illustrative of them are the following which are selected from approximately two dozen such opening sentences of paragraphs in Chapter III.

Every organism strives to satisfy its needs by seeking appropriate relationships with a variety of preferred environmental conditions.

Pupils differ in all components of the psychological adjustment process.

There are tremendous differences among pupils in their responses to chronic or prolonged frustration.

Numerous well-worded generalizations do not, of course, guarantee a book's worth. The principal criterion is the usefulness which the generalizations have for the reader—in this case, a prospective teacher. To be useful to teachers the ideas and principles presented in a textbook on educational psychology must be soundly based on research, be accurately interpreted, and of significance to teachers.

In the main, the content of *Educational Psychology* measures up to these demands. For the most part, the conclusions which have been presented are founded in research. The references cited at the end of the chapters show the author's familarity with the best that research in the behavioral sciences has to offer. Not everyone will agree entirely with their choice of the research findings and principles around which to develop such a textbook, but they have clearly made a determined effort to select from psychology that which has significance for teachers and to avoid the inclusion of topics of little real value to teachers.

One topic which the authors did not altogether avoid but to which they gave surprisingly little attention is that of perception. Possibly they felt that they had done justice to it in their chapter on concept growth where they touched briefly on

certain of the effects of the way the individual perceives himself and his world. Nevertheless, for perception to be limited to two or three brief mentions is difficult to understand in the light of the extensive current interest in the topic on the part of psychologists and educators.

The book is profusely illustrated. Not less than the equivalent of 40 pages is taken up by illustrations (drawings and cartoons) and not less than 30 pages by graphic and tabular figures. The number and variety of the illustrations (approximately 100) are disconcerting—especially those whose relevance to the discussion is not apparent. The tables and graphs, however, are uniform in size and are attractively presented—some have a pleasing touch of color.

There is another kind of illustration which the authors of this book have employed to much better advantage. They have made considerable use of case studies and anecdotes. Many of the principles which are introduced in the book will be more easily grasped by the reader because of the help he gets from these well-chosen case materials.

University of Nebraska WARREN B. BALLER

Chemical Concepts of Psychosis. Proceedings of the Symposium on Chemical Concepts of Psychosis held at the Second International Congress of Psychology in Zurich, Switzerland, Sept. 1 to 7, 1957. Edited by MAX RINKEL with HERMAN C. B. DENBER. New York, McDowell, Obolensky, 1958. Pp. xxi, 485. $7.50.

Reticular Formation of the Brain. Edited by H. H. JASPER, L.D. PROCTOR, R. S. KNIGHTON, W. C. NOSHAY, and R. T. COSTELLO. Boston, Little, Brown and Company, 1958. Pp. xiv, 766. $16.00.

Neuropsychopharmacology. Proceedings of the First International Congress of Neuropsychopharmacology, Rome, September, 1958. Edited by P. B. BRADLEY, P. DENIKER, and C. RADOUCO-THOMAS. Amsterdam, Elsevier Publishing Company, 1959. Pp. xv, 727. $27.00.

These three recent international symposia continue the story of the striking developments in behavioral physiology within the last ten years. They differ considerably in their format. *Chemical Concepts of Psychosis,* the most fully integrated volume of the three, is one which, because of this integration, will probably retain its reference value longest. The arrangers have thought to preface the research reports with historical papers by Osmund and by Baruk, and have then arranged the papers in a way which ought to be a model for future symposia. Part II, comprising the bulk of the volume, contains discussions of *LSD,* mescaline, epinephrine, serotonin, and the indoles. Each drug is usually represented by a paper on its chemistry or pharmacology plus one or more papers on psychological and clinical effects. Papers follow on psychological effects of extracts of schizophrenic blood, on neurophysiological studies in animals, and on problems relevant to drug studies. Tranquilizers are discussed in Part II which is followed by four papers on the relationship of these new ideas to the disease entity, schizophrenia. Four papers in Part V comprise "Reflections" with a concluding discussion. The volume will not substitute for an authoritative summary of these new developments, but since no such account is as yet available, the arrangers are to be thoroughly congratulated on producing a research-oriented book of great value to others than research workers in the field. This book can be wholeheartedly recommended to interested psychologists as the best introduction to psychotropic drugs yet written in symposium form.

Reticular Formation of the Brain contains 36 papers and a summary from the 1957 Henry Ford Hospital symposium. Designed as a sequel to the now-famous symposium on *Brain Mechanisms and Consciousness* (1955), this volume has already acquired standing of its own. Each chapter is topical and few reviewers are included, but for persons with some background in this area, there are high-level chapters on the neurophysiology of learning by Ingram and by Lilly (subcortical stimulation); by Morrell and by Gastaut (conditioned changes in the *EEG*); on self-stimulation experiments by Olds and by Brady, on perception by Lindsley; and on hormonal changes induced by physical and behavioral stressor stimuli by Mason. For psychologists who wish to further their understanding of the reticular formation, the arrangement of chapters will give little aid, but Papez' chapter is a good account of anatomical pathways. Physiological anatomy is well represented by French (descending pathways), by Green and by Adey (relationships with the rhinencephalon), and by Ward (motor effects). New information on the striking sensitivity of the reticular core to drugs is presented by the Killams, Bradley, Domino, and Dell. Chapters by Harris, Sawyer, and Hume describe some pathways of hormonal activation.

If anyone doubted that a strict distinction between neuroanatomy and biochemiscal physiology can no longer be maintained, then the coinage *Neuropsychopharmacology* should almost by itself convince him. These preceedings contain seven invited addresses, several which will be of interest to psychologists working with animals (contributions by Bradley, Olds, and Masserman) or with clinical patients (Hoff and Arnold, Mayer-Gross, Bleuler, and Elkes). Clinical psychologists will also be interested in two plenary sessions devoted to the impact of psychotropic drugs on psychiatric hospitals. Following these are 132 brief research communications, some already published, but of such widespread coverage that all psychologists in this area will find something of interest (the writer found 22 of interest to him, and others will undoubtedly find more). One point of closure seems close at hand. Olds presents evidence that positive reinforcing points in the rat brain are identical with parasympathetic areas and negative reinforcing with sympathetic areas of the midline structures. In a thoughtful discussion, Kaada (pp. 38-45) supports these claims. Olds also finds that the positive and negative centers respond differently to psychotropic drugs and from this proposes a theory of psychosis. Intriguing though this suggestion is, several contributors urged caution for speculations of this kind.

Although symposia may sometimes seem to be multiplying according to a law all their own, these three volumes document a steady, difficult progress toward elucidating the physiological mechanisms of motivation, learning and psychosis. The often-promised breakthroughs have yet to materialize, but those of us who find the hope for breakthroughs to be illusiory can nevertheless be cheered by the certain modest but positive advance shown in the papers of these volumes.

Cornell University ALLAN C. GOLDSTEIN

Design for a Brain. By W. ROSS ASHBY. London, Chapman & Hall, 1960. Pp. 285. 42s.

There can be no scientific issue concerning the concept of brain as machine in Ashby's view. All that is meant by a "machine" is that it is a state-determined system, and all scientists operate on the assumption that state-determined systems can be found for every phenomenon they study. Scientific issues must therefore concern

the nature of the machine. Ashby's design—now revised to a major degree from the orginal version published in 1952—is an hypothesis about this nature.

What are the fewest restrictions to the specifications for physical systems in general that will permit the selection from a universe of systems those which behave like cerebro-environmental systems? This is the problem the author is primarily concerned with. To solve it he specifies wide classes of system, notes what effects would follow from different, structurally minor variations of the more general types, and compares these effects with what is known about the behavior of bio-environmental and cerebro-environmental systems. This enables him to specify design characteristics for classes of system successively one stage less general and one stage closer to the characteristics of cerebro-environmental systems.

Three notable varieties of system are examined in the progress of Ashby's argument. The first, the system with feedback, is capable of producing simple adaptation or coördination. The simplest sorts of adaptive mechanism (showing 'lines' or 'trajectories' of action) are at least this complex. Such is the 'stable' system. The behavior of living systems may be even more complex however.

Certain variables must be kept within specified limits for the maintenance of life. Insofar as behavior displays coördination (within a 'field') with reference to these essential variables in addition to coördination along any single trajectory, it gives evidence of secondary feedback. This is the ultra-stable system. The electro-mechanical 'homeostat' designed by the author has the characteristics of such a system.

The specifications for ultra-stable systems are not yet sufficient for the selection from a universe of systems those which will act like cerebro-environmental systems. Several cogent reasons for this conclusion are presented. In particular such systems are too slow in adaptation, and they have no potentiality for repetition of previously adaptive lines of behavior upon repetition of a disturbance. For both of these reasons an additional specification is required to produce considerable independence of parts in the system. Any such specification, however, must be compatible with the primary requirement of all systems for interdependence of parts. The author's general solution is to specify that the system be constructed of parts that have a high proportion of their possible states as states of equilibrium. This will produce local constancies which will split the whole in a fluctuating and partial way. Such 'polystable' systems are proposed to approximate cerebro-environmental systems to such a degree that the latter must be presumed to belong to this class.

The application of systematic terminology to bio-environmental relations demands as much attention to the environment as to the biological parts. Ashby supplies this, and his discussion of environments is both enlightening and stimulating. Here, as indeed throughout this book, Ashby's terminology has the distinct advantage that it frees one from the improperly simplified conceptions of what is going on when an animal adapts or fails to adapt that are entailed in other manners of speaking.

Granting the ultimate benefit of "other categories" for use in thinking, in what other way can Ashby's contribution assist the student of behavior? There are two answers: first, by simplying the subject matter to the greatest degree consistent with exact and durable conclusions; and secondly by clearing up epistemological and analytical confusion. The contention that all science is Cybernetics pervades Ashby's thinking, both in this book and in its companion (*An Introduction to Cybernetics*, 1958). Thus an experiment is a system with feedback and what is fed back is information in the technical sense of communications theory.

The book is written in admirably lucid English. Its assertions are at once philosophically general and mathematically precise. Its differences from other essays are not superficial but fundamental. Its arguments and viewpoints will interest all students of complex systems for years to come.

Borden Neurological Institute
Bristol, England
J. A. EASTERBROOK

Toward Understanding Human Personalities. By ROBERT W. LEEPER and PETER MADISON. New York, Appleton-Century-Crofts, Inc., 1959. Pp. xiv, 438. $5.50.

The conceptual openmindedness and consequent broadening of experimental investigation which has developed with the decline of general theories is nowhere more apparent that in the study of personality. While there can be little doubt of the amount and extent of research and theoretical activity in this area, the questions of significance and fruitfulness are persistently raised. This book illustrates many of the difficulties, stresses and strains, conciliations and new directions which have developed within personality research. The answer to the question of whether the present integration is successful, like the question of whether the field is making progress, will clearly depend upon the criteria of the reader. Those who insist upon rigor and verification, and those who measure progress in terms of major changes will be disappointed. Those more moderate in their demands on these dimensions will find this an involving and insightful book.

The approach of Leeper and Madison is based largely upon an integration of the theories of the 'dynamic self' derived from psychotherapy (Rank, Sullivan, Rogers, etc.) with those aspects of the Gestalt and purposive theories which emphasize complex organizing processes (Köhler, Lewin, Tolman, etc.). The authors propose an organizational theory of personality which emphasizes perceptual and conceptual processes. Perception is extended to include "processes that are highly motivational, that are conceptual in character, that sometimes are unconscious and that frequently are long sustained" (p. 194). The nature of these processes is, to a great extent, determined by past learning. Individuals develop characteristic perceptual habits by means of which previous psychological realities are reintegrated with those immediately present. Personality is largely a matter of these perceptual habits. The customary distinctions between conscious and unconscious, and perception, abstraction and learning are thus dissolved and the product is integrated with the concept of a field which has a motivated self at its center. The applications of the theory are extended to such topics as personality development, adjustment (unity and disunity), transference (reintegration), and therapy (the changing of personality).

Although the book is intended primarily for laymen and students, the authors state their hope that it will make a theoretical contribution. Their attempt to reach this broad audience is apparently the reason for the virtual independence of its three parts. The first part deals with the various meanings of personality. The second examines psychotherapy, anthropological data and experimental psychology as sources for understanding personality. These first two parts are clearly intended for the general reader whereas the third section is the "meat" of the book and apparently its reason for being. It is this third section, in which the organizational theory of personality is developed, which will be of interest to the professional reader. It is opinion of this reviewer that in attempting to reach such a diverse audience the

authors have detracted from the genuine merits of their book. The general language and level of complexity of an introductory treatment of personality are not ideal for the development of theoretical advances.

The many years of preparation show in the clear and concise writing and the well developed and integrated ideas. An interesting, smooth flowing style effectively conveys a sense of understanding and the feel of human experience. The organization of the material, both in terms of ideas and physical appearance, is unusually well done and attractive. In their emphasis upon developing a theoretical position, the authors do not indicate the bases for their generalizations. Only rarely are experiments described and much of the evidence is based upon individual cases or the personal experiences of the authors. While espousing the need for research, the book does not indicate how its concepts are testable, or indeed, how research proceeds.

Optimistic and positive in its orientation, well-organized and presented, the book will probably be most effective with introductory courses of personality. More advanced students would benefit by the many provocative insights contained in the third section. Like Diamond's *Personality and Temperament* of two years earlier, it indicates a reawakened interest in self psychology and inherited psychological characteristics. This book can be criticized, but it is too stimulating and sound to be passed over. Students who use it will be more aware of the richness and organization of human behavior than is generally the case.

Wesleyan University — STANLEY COOPERSMITH

Right-Left Discrimination and Finger Localization: Development and Pathology. By ARTHUR L. BENTON. New York, Paul B. Hoeber, Inc., 1959. Pp. xv, 185.

The alterations of behavior produced by cerebral disease in man are seldom studied by experimental methods. Physicians are usually not equipped by training or orientation to employ such methods in the study of these disorders; psychologists, who are equipped, have perhaps not generally recognized that the forms and patterns of functional disintegration in cerebral disease can provide valuable clues to normal psychological organization. Professor Benton's monograph is an example of systematic investigative methods applied to certain behavioral disorders which have excited interest and controversy for many decades. His principal aims were to develop reliable indices of ability, to establish norms against which the performance of mental defectives and brain-damaged subjects could be evaluated, and to describe the results of administering his tests to such deviant groups. In a field where formal tests of behavior are still not quite *comme il faut,* and where the use of a control group is a rarity, these objectives deserve mention.

Impairment of lateral localization of body parts (right-left disorientation) and loss of ability to identify individual fingers on either hand (finger agnosia) are two of the four deficits constituting Gerstmann's syndrome, the other two being agraphia and acalculia. These deficits are well worth investigating for several reasons. According to Gerstmann, these complex disturbances occur as a unit and are dissociated from other deficits; he attributes their concomitance to the close psychological relationship in development of the functions disturbed. Furthermore, he states that the syndrome results from unilateral lesion of a small area of the left cerebral hemisphere. These relationships, if established, would obviously have important implica-

tions for physiological and developmental psychology, as well as for several applied disciplines.

Professor Benton's work, combined with his thoughtful appraisal of the literature, casts serious doubt on these claims. In the first place, neither right-left discrimination nor finger localization is in itself a unitary capacity. Different methods of assessment show that various aspects of each ability do not necessarily stand or fall together. Writing and calculation likewise appear to have numerous determinants. Secondly, although right-left discrimination and finger localization show a moderate degree of relationship, there is no invariable association of deficits in these abilities. A point of special interest is that the dependence of arithmetical learning on finger differentiation, as postulated by Gerstmann, is not supported empirically. Finally, Professor Benton's studies indicate the desirability of considering right-left discrimination and finger localization within the broad context of general mental level and especially of conceptual-symbolic functions. In children, the relationship to mental age is marked. Retardation in language development appears to play a significant role in certain types of deficit. In brain-damaged patients, mental deterioration, dysphasic tendencies, and spatial disorders of various kinds frequently accompany, and may perhaps determine, right-left disorientation and finger agnosia.

The experimental studies reported here are more complete with respect to analysis of developmental trends than of acquired impairments. More extensive investigations of brain-injured adults, and a wider sampling of functions impaired and preserved in these subjects, might have yielded direct experimental evidence relevant to Gerstmann's claims and to Benton's own hypotheses. His evaluation of the concept of the body scheme would likewise have gained by empirical investigation of whether or not disorientation affecting the body and disorientation in external space are actually independent. Further work on specific interrelations among children's abilities would also be desirable; since the mental age is derived from an averaging of abilities, this measure should be supplemented by others capable of revealing particular functional dependencies. The studies concerned with language development and with arithmetical ability, previously mentioned, represent a beginning of this type of behavioral analysis. Another area in which experimentation might profitably be extended is that of cerebral mechanisms in the behavior studied. Brain-behavior relationships are not a central theme of this monograph; in this reviewer's opinion, however, the discussion would have been strengthened by a more skeptical attitude even than Benton's toward notions prevalent in clinical neurology about the nature of functional representation.

New York University JOSEPHINE SEMMES
Bellevue Medical Center

The Motivation to Work. By FREDERICK HERZBERG, BERNARD MAUSNER, and BARBARA B. SNYDERMAN. New York, John Wiley and Sons, 1959. Pp. xv, 157, $4.50.

Whether he is a practitioner of human relations or a motivational theorist, anyone seriously interested in the motivation of the normal human adult should read this book. There are implications here for research, and even for immediate application, despite serious limitations of methods of research, reporting, and statistical analysis. A summary of the procedures of these investigators will highlight these limitations.

It may also be useful to the reader, since the authors' exposition is not always clear.

Nearly 200 (reviewer's estimate) engineers and accountants in nine plants in the Pittsburgh area were interviewed to obtain reports of sequences of events during which they had felt unusually favorable or unfavorable about a job they had held. Each gave at least two sequences (mean, 2.4) including one short-range sequence (an event comparable to a critical incident) and one long-range sequence (a series of events lasting over a considerable but finite period of time). One of these sequences was favorable ('high') and one unfavorable ('low'). After elimination of some of the sequences, the 476 remaining were analyzed for content, using an *a posteriori* code.

Many comparisons ('high' vs. 'low,' long vs. short-range, engineers vs. accountants, 'high' and long-range vs. 'low' and long-range, etc.) were made for the frequencies in which different themes were mentioned, significance being tested presumably by *t-test.* (There was no cross-validation group).

During the same interview, the relation of the job feeling to performance, turnover, mental health, interpersonal relations, and attitudinal effects were retrospectively reported. Analysis was similar to that for the sequences.

The problems of generalizing from such a study to other situations, individuals, jobs, occasions, and methods of investigation are all too evident. Most of the limitations of the study have been explicitly or implicitly acknowledged by the investigators themselves. Nevertheless, the obtained differences in some instances are so large and so consistent from comparison to comparison that they are quite convincing.

The most generally interesting finding is that more different factors operate in the 'high' than in the 'low' sequences. Achievement, recognition, the work itself, responsibility and advancement appear more frequently for the favorable sequences, while interpersonal relations, supervision, company policy, working conditions, and personal life appear more frequently in the unfavorable. Related to these differences, improvement in performance is reported more frequently with the 'high' sequences than is a reduction in performance in the 'low.' The investigators relate the 'high' factors to 'motivators,' which facilitate production and output beyond the bare minimum of the accepted 'fair day's work,' and the 'low' factors to hygiene which prevents active dissatisfaction and may, at best, reduce restrictive practices, but will not produce satisfying involvement with the job. The distinction between 'satisfiers' and 'dissatisfiers' and their relations to actual overt behavior should certainly be investigated further not only with simple comparisons of frequences, but with correlational and experimental procedures, and in the framework of a more careful system for anchoring the scales of judgment used.

The publication of these findings as a book rather than as a journal article or monograph can only be justified by the quality of discussion of the general implications of the results. Despite the difficulty which the reader experiences in disentangling the usage of such undefined concepts as satisfaction, involvement in job, attitudes, factors, and feelings, he will find relationships and comparisons which are interesting for theory. The personnel administrator now 'applying' unvalidated theories of motivation at work will also find a challenge to the assumption underlying his policy as well as possibilities for real improvement in its effectiveness.

Cornell University

PATRICIA C. SMITH

Elementary Statistical Methods in Psychology and Education. By PAUL BLOMMERS and E. F. LINDQUIST. Boston, Houghton Mifflin, 1960. Pp. xv, 528. $5.75.

Information and Error By SOLOMON DIAMOND. New York, Basic Books, 1959. Pp. xii, 307. $5.00.

Basic Statistical Methods. By N. M. DOWNIE and R. W. HEATH. New York, Harper & Brothers, 1959. Pp. xii, 289. $4.50.

Blommers and Lindquist have written a really excellent book which stresses basic statistical reasoning. The text is well unified, the examples are good, and it is evident that an immense amount of thought and care have gone into the exposition and the choice of topics to be covered. Moreover, the book is free of nearly all the usual confusions and misconceptions.

Because of the emphasis on fundamental ideas and the care spent in developing them, the book, despite its length, does not cover all the topics usually found in introductory texts. The first eight chapters are on descriptive statistics. The next three provide a 100 page introduction to statistical inference: sampling, hypothesis testing, and confidence intervals. The chapter on hypothesis testing is magnificent. It should be considered for outside reading in more advanced courses. The book concludes with a chapter on *t*, and three fairly long chapters on correlation and regression.

The instructor considering the use of Blommers and Lindquist's text is warned that it is not easy reading. In large part, this is because a fair amount of honest thought is required of the reader. A part of the difficulty also stems from the authors' style which tends to be somewhat ponderous. Despite this drawback, the book will set a standard of excellence for some time to come.

Diamond's *Information and Error* has two principle characteristics. The first is the novel organization which is centered on the notion of variance. This organization produces a considerable economy of presentation. It also yields a treatment which is noticeably more relevant to experimental psychology than that which results from the orientation toward individual differences which pervades most introductory texts. Noteworthy minor features include slide rule instruction, a very sensible chapter on the analysis of nonnormal data, and a chapter on factor analysis. Despite the title, the book has nothing to do with information-theory.

The second main characteristic of Diamond's book is its breezy, graphic style. Although the reviewer thought that the net effect was considerably to the good, this feature, together with the novel oraginzation, may make the book unpalatable to some instructors. In addition, although the treatment is statistically competent, the development of a number of topics is rather casual and will probably leave many students without adequate anchorage unless supplemented by lectures.

Downie and Heath's *Basic Statistical Methods* covers a standard array of topics for an introductory course. In addition, it devotes a chapter to test theory, and to distribution-free statistics. Every chapter is provided with a set of exercises. The several initial chapters on descriptive statistics are very readable. There is an emphasis on computation which is integrated with the text with surprising smoothness. Unfortunately, the treatment of significance tests is less well done. Computational procedures are sometimes cumbersome. The discussion of statistical inference is quite

poor, *e.g.* "At the 5% level, there are 5 chances in 100 that we are wrong, that is, that the hypothesis we are rejecting should not be rejected. The 1% level cuts these chances of error to 1 in 100" (p. 120). Every statistics book should be allowed a certain number of slips and minor pedagogical inaccuracies. This quotation, however, is neither of these but exemplifies the treatment of statistical inference.

University of California, Los Angeles NORMAN H. ANDERSON

Der Labyrinth-Test. By FRÉDY CHAPUIS. Hans Huber, Bern, 1959. (Two editions, German and French.) Pp. 147. DM 24.50.

After reviewing the history of the labyrinth-test (*i.e.* the maze-test as it has been applied to children and adults, Chapuis rejects the test as being too easy for personnel research and diagnosis. He describes a much more difficult procedure which he has applied to 5 groups of subjects: boys of 16–20 yr. of age; employees in retail sales and the like from 19–30 yr. of age; semi-skilled factory workers; skilled workers and technicians (such as locomotive engineers and air pilots); and untrained laborers.

Every *S* taking the test is given a red pencil as a stylus and three mazes, through which he is to trace a path. The three mazes are printed upon a single sheet of paper 30 cm. high. The smallest and easiest maze is at the top, the largest and most difficult is at the bottom, with the maze intermediate in difficulty between the other two. *S* is told not to waste any time, but there is no time limit to the test. At a signal, the test-sheet is turned up and *S* is told to solve all three problems, beginning with the smallest and ending with the largest.

The tester watches *S* and, on the basis of a scheme which he has memorized or has before him, seconds, in half- and whole-unit errors, the mistakes made by without telling him what mistakes he has made. In addition, a record is kept of the total time required by *S* to complete the three mazes. After the test has been completed, the tester writes down what he has observed regarding *S*'s general 'characteristics,' *i.e.* his activity during the test, the pressure he placed upon the pencil, and the regularity of his movements; the degree of his confidence and the exactness of his work.

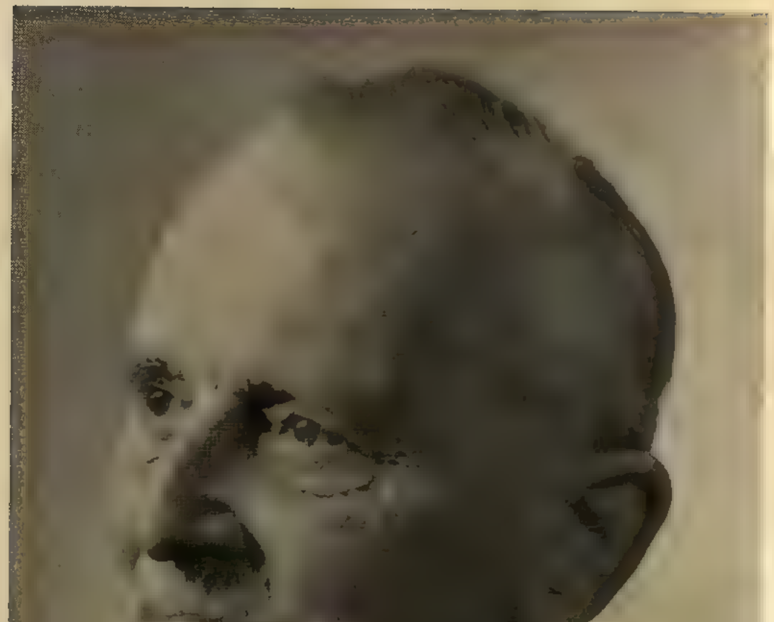
The *S*s differ greatly. Some show an exaggerated caution; some 'seem to lose themselves completely'; some are impulsive and impatient, and some have no regard for exactness and work as though they do not understand the directions given them. A few are regular 'swindlers,' who use the tester's moments of seeming inattention to break the rules of the test.

Chapuis has worked out a table in which the number of errors and the time required to complete the mazes are translated into a quantitative score.

The test has been used officially in Switzerland. An amusing example is the case of a drunken auto driver who insisted *in court* on his freedom because he was not, as he claimed, affected by liquor. The test was given him in the morning and repeated in the afternoon, that is, before and after the consumption of liquor. The results clearly showed that alcohol was effective.

The book gives 46 photographic illustrations of what the performers actually did with their red pencils.

Miami, Florida MAX F. MEYER



(See page 644)

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LXXIII DECEMBER, 1960 No. 4

SPATIAL LOCALIZATION OF AFTER-IMAGES

By J. L. ZAJĄC, Unversity of Edinburgh, Scotland

Investigations on after-images were carried out by all of the principal investigators in the field of visual perception. Numerous theories of after-images were advanced which were used to support or to criticize one or another theory of visual perception. Fechner, Plateau, and Helmholtz, each advanced his own theory.[1] As concerns the position of the images, they were mostly located at the distance of the background on which they appeared, and they were considered as two-dimensional images or colored shadows attached to the background. When the eyes of *O* were closed or when the background was dark, the images appeared to wander freely about the visual field. Some *O*s had the impression that when seen as smaller they were perceived farther away.

Hering contended that the size of an after-image depended on the distance of the background on which it was perceived and not on the distance of a fixation-point placed between the *O* and that background.[2] This contention, which was intended to overthrow the projectionists' theories of space perception was, and is still, so suggestive that it is repeated in the most recent publication on the binocular perception of depth by Linschoten.[3]

* Received for publication July 29, 1959. My thanks are due Professor James Drever for his assistance in this investigation; to the University Court and the Senatus Academicus of the University of Edinburgh for grants-in-aid which made this study possible; and to Messrs. Robert Mulvey, G. F. R. Henderson, and R. S. Corteen for serving as *O*s.

[1] G. Th. Fechner, Über die subjektiven Nachbilder und Nebenbilder, *Poggen. Ann.* 50, 1840, 193-221, 427-465; Joseph Plateau, Über das Phänomen der zufälligen Farben, *Poggen. Ann.* 32, 1834, 543 ff.; *Essai d'une Théorie générale comprenant l'ensemble des apparences visuelles qui succèdent à la contemplation des objets colorés*, 1834; Hermann von Helmholtz, *Treatise on Physiological Optics*, Trans. 3rd ed., J. P. C. Southall, ed., 2, 1924, 228-262.

[2] Ewald Hering, *Beiträge zur Physiologie*, 2, 1862, 132 ff., 150.

[3] Johannes Linschoten, *Strukturanalyse der binokularen Tiefenwahrnehmung: Eine experimentelle Untersuchung*, 1956, 49 ff.

Several authors dealt with the problem of relations between the size of the after-images and the geometrical distance of the background on which they were seen. The best formulation of this problem was given perhaps by Emmert.[4] The results of his investigations and his conclusions, published after Hering had advanced his thesis, deal only with the relation between the size of the after-images and the distance of the background on which they are seen. Emmert nowhere mentions the relation between the size of after-images and the distance of the fixation-point. The so-called Emmert's law tells us that there is direct proportionality between the size of after-images and the distance of the background. In his investigations he does not discriminate between the perceived and the geometrical (actual) size of the after-image, and these are two different things. Nor is he concerned with the localization of after-images in the third dimension.

The investigators concerned with stereoscopic vision noticed a relief in their after-images, *i.e.* differences of distance of different parts of the after-images from them. Wheatstone, Rogers, and Wundt obtained relief of after-images produced by fusing monocular after-images, which were formed at not exactly corresponding points of the two retinas.[5] Wheatstone concluded that disparate after-images can produce stereoscopic depth. Rogers found that he could develop first an after-image in one eye, then another in the other eye and finally combine them stereoscopically. Helmholtz obtained distinct apperception of depth from positive after-images which he developed by gazing for a moment at some brilliantly lighted object.[6] Washburn and Smith found with 19 *Os* the existence of relief in after-images.[7] Ferree and Rand, Fernberger, and Wohlzogen were other investigators who have experienced relief in after-images.[8] Several statements have been made, then, concerning the corporeity of after-images, but no direct investigation has been made of their depth.

Problem

The present investigation was undertaken to discover whether the perception of depth could be obtained from after-images and, if so, to compare it with the perception of depth from objects and stereoscpic images.

Method of Procedure

Four persons served as *Os*. Objects in the form of colored vertical rods 3, 6, and

---

[4] Erasmus Darwin, On the ocular spectra of light and colours, *Phil. Trans.*, 76, 1786, 313; E. Emmert, Grössenverhältnisse der Nachbilder, *Klin. Monatshäfte Augenhk.* 19, 1881, 443-480.

[5] Charles Wheatstone, Sur un effet singulier de juxtaposition de certaines couléurs dans les circonstances particulières, 1845, 75 ff.; W. B. Rogers, Observations on binocular vision, *Silliman's Journ.*, 31, 1856, 80-95, 173-189, 439; Wilhelm Wundt, *Beiträge zur Therie der Sinneswahrnehmung*, 1862, 286-287.

[6] Helmoltz, *op. cit.*, 3, 1925, 456.

[7] M. F. Washburn and D. L. Smith, Stereoscopic binocular fusion in the original impression and in the negative after-image, this JOURNAL, 45, 1933, 320-321.

[8] C. E. Ferree and Gertrude Rand, Perception of depth in the after-image, this JOURNAL, 46, 1934, 329-332; S. A. Fernberger, A figural after-effect in the third dimension of visual space, *ibid.*, 61, 1948, 291-293; F. X. Wohlzogen, Die Entstehung sterischer Nachbilder, *Experientia*, 8, 1952, 194.

8 mm. in diameter were used as primary stimuli to produce after-images. They were placed on the bars of Zając's Natural Stereoscope.[9]

Red proved to be the color best suited for the experiments. For producing after-images, fixation lasted about 20–30 sec. When changes resulting from prolonged fixation were to be observed, longer intervals were used. For fixation upon primary objects, a 100-w. electric bulb was placed in a dark room 4 ft. above the apparatus; for observing after-images this light was turned off and *O* looked at the floor which was covered with dark linoleum, illuminated only by a 100-w. lamp placed about 15 ft. away. As a new fixation-object another metal rod with pointed top was used. It was held stationary by *O* in his line of regard or moved slowly back and forth from a position near his viewing eye to one at arm's length away.

The experiments on after-images were conducted in 1955-1957. They were divided into two groups: (1) those with monocular vision; and (2) those with binocular vision. The experiments with binocular vision were subdivided into four groups: (a) with double images—one rod used for fixation and a second rod was seen in double images, both being placed in the median plane; (b) after-images produced by fixation upon one rod and of another rod or rods placed laterally near the horopter and seen singly; (c) after-images produced by stereoscopic images of so-called 'walls' to be described below; and (d) after-images produced by 'walls' in special form of straight or inverted 'V' with a small interspace in the middle.

Results

The first important finding was that, independently of background, after-images can be caught, like fish on fishrods, on fixation-points. If we stabilize a fixation-point upon an object in front of us, then the after-image will join it, will come and stay at its distance or slightly behind. It is not, however, very steady, it is rather 'trying' to fall out to the right or left, or back, up or down and it is necessary to make some effort to hold it at the fixation-point. This is, of course, due to eye-movements, to the fact that the eyes do not hold a steady, fixed position, but tend to move about. This difficulty was also mentioned by Helmholtz. In a stabilized position the part of the after-image representing the fixation-point in the primary object is seen as equidistant from *O* with the new fixation-point, and then its size, both perceived and geometrical, is related to the distance of the new fixation-point. When we move the fixation-point, the after-image moves slowly with it and it grows in size with increasing geometrical distance of the fixation-point and diminishes in size with decreasing distance from *O*. This is the reverse of what happens to the size of the fixation-point itself. These phenomena are in contradiction with the observations of Hering's mentioned above. They also mean, that the

---

[9] J. L. Zając, Depth perception of double images in the vicinity of other images. *Acta Psychol.*, 12, 1956, 111-129.

so-called Emmert's law must be interpreted in the sense that the geometrical size of after-images is proportional to the geometrical distance of the fixation-point both in monocular and binocular vision. As concerns depth-perception of after-images, their localization is relative to that of the fixation-point. The absolute distance of the latter from *O* is not an object of perception but rather of estimation.

Striking observations can also be made in the following way. We first look for 20–30 sec. at a colored square or circle, as on a primary object. We then throw the negative after-image at a marked fixation-point on a wall. The geometrical size of this image will be related to the difference of distance from *O* between the primary object and the wall. The corresponding changes will occur also in perceived sizes. Afterwards, always fixating the marked point, we approach the wall: we observe then a continuous shrinking of the after-image on the wall. At the nearest distance from the wall (within the range of accommodation of the eyes), the after-image becomes very small. That shrinking can be observed with one or with both eyes. The other observation is as follows. If one happens to be able to effect convergence of the two eyes beyond the wall, which can easily be done by an experienced *O*, then the after-image will appear to grow in size; this increase in size is accompanied by an impression that the image is floating somewhere behind the surface of the wall. We ourselves can easily ascertain whether this convergence is really made beyond the wall, because in that case the point on the wall, seen singly when fixated, will be seen in crossed double images. When convergence is effected in space between the wall and *O*, then the after-image will become smaller and will be seen floating somewhere between the wall and *O*. In the latter case, the point on the wall will be seen in uncrossed double images. It is not so easy with monocular vision, where fixating a point within the range of accommodation means accommodating the eye on it and it is difficult to accommodate the eye at an unreal object beyond the wall. One can, however, bypass that difficulty in the following way. One first looks for 20–30 sec. at a colored object with one eye (the other eye being closed), thus producing an after-image in one eye only. One then throws the after-image on the wall, approaching it very closely and fixating steadily the marked point on it: thus diminishing the size of the after-image. Now the other eye is opened and the convergence beyond the wall is effected; it is now certain that both eyes, as well as each eye separately, are fixating a point beyond the wall. Then the after-image is seen (only with one eye) as greater than when the point on the wall was properly fixated. In this case, however, the after-image which was seen distinctly with one eye, is now seen with both eyes somewhat diluted (contrast effects and other factors are here involved).

Similar phenomena are observed when we throw an after-image on a surface reflected in a mirror: this image looks larger than an image seen at a distance of the surface of the mirror itself. These phenomena must also be regarded as being in contradiction with Hering's findings.

The change in the apparent size and distance of after-images might be compared and related to other phenomena of space-perception de-

scribed by the author.[10] It has been found that when the distance of the fixation-point is changed, the objects in the lateral field of view grow and come nearer as the fixation-point is moved away from *O*, and diminish in size and move away from him as the fixation-point comes nearer to him. These facts can be observed in monocular as well as in binocular vision with double and stereoscopic images.

In the case of real objects in the field of vision, the changes in size of their images are due both to the law of visual angle and to the changes in accommodation. Within the range of accommodation the curvature of the crystalline lens diminishes with growing distance of the fixation-point and the retinal image grows with it. The law of visual angle has two aspects. (1) The angle of an object in the field of vision subtended at the eye diminishes with the growing distance of that object, when fixated. (2) The angle of another non-fixated, stationary object remains unchanged, its retinal-image size remains also unchanged (except for the changes due to accommodation) but its linear geometrical size and lateral distance from the fixated object grow with the growing distance of the fixation-point, and with it grows also the perceived size of the non-fixated object (see Fig. 1).

(1) It is well known that in monocular vision for the near field of view the localization in the third dimension is different from that obtained with binocular vision. The monocular depth perception is not so certain and steady. When fixation is of short duration, one can perceive differences in distance from *O* between the objects placed laterally and the fixation-point. Soon, however, there appears a tendency for these differences to be diminished, changed, or effaced; images of objects in the lateral field of view approach the plane of the fixation-point. Sometimes so-called inversions occur: farther objects are seen nearer than the fixation-point or nearer objects are seen farther away. In other cases, the *O*s state that it is very difficult for them to tell by monocular vision the relative perceived distance of objects placed laterally in comparison with the fixated object.

For producing after-images in our case it was necessary to prolong fixation for about 20–30 sec., so it was difficult not to observe changes occurring in perceived distances mentioned above. As the result of the experiments on after-images with monocular vision it has been shown that generally all the changes observed by prolonged fixation on primary objects occur also in after-images, but very often the after-images of objects placed

[10] Zając, Investigations on depth perception in indirect vision, *Brit. J. Psychol.*, 44, 1953, 132-144.

laterally are seen at the same distance as the after-images of the fixation-objects, or their distance relative to the distance of the after-image of the fixation-object is very difficult to determine.

(*2 a*) There are several contradictory theories about the depth-perception of double images. Some authors state that they are seen on the surface of the frontal plane passing through the fixation-point; others, that they are perceived on the hemispheres passing through it; still others, that they are on the visual axes—all these belong to the so-called projectionists' school of space-perception of different types (Aquilonus, Nagel, Schulz).[11]

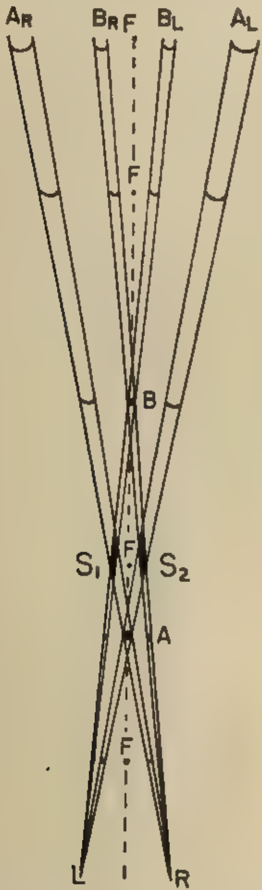


FIG. 1. DIAGRAM ILLUSTRATING THE LAW OF VISUAL ANGLE

The greater the distance of the fixation-point from *O*, the greater the size of images of objects placed in lateral field of view and also the greater their lateral distance from the fixation-point. L = left eye, R = right eye, A,B, = the two vertical rods; S_1, S_2 = stereoscopic images; F = fixation-points; the semi-circles denote the sizes and lateral distances of monocular and double images but not perceived distances from *O*, which might be quite different.

[11] Aquilonus, *Opticorum Libri VI*, Antwerpiae, 1613; A. Nagel, *Das Sehen mit zwei Augen und die Lehre von den identischen Netzhautstellen*, 1861, 1-184; Schulz, Ueber physiologische Gesichts- und Farben-Erscheinungen, *Schweigerts' J. physik Chemie*, 1816.

And for still others they are seen at the distance of the objects seen in double images when these objects are fixated (Helmholtz, Hering, Aall, Tschermak).[12]

In reality, double images are seen at various distances from *O* and their positions often change. It depends on different factors, among others, on the lateral distance from the fixation-point. Our results show that the nearer the double images are seen laterally to the fixation-point, the more they are subject to the laws of stereoscopic depth-perception; and that the farther they are apart from the fixation-point, the more their position becomes indetermined, or the more frequently they are seen at the same distance as the fixation-point.[13] With fixations of short duration, the double images are seen somewhere between the fixation-point and the object seen in double images when it is fixated. Their position depends also on the duration of the fixation—the longer the fixation, the more frequently do changes in depth occur in the perceived position of double images. Very frequently so-called inversions occur which consist in the reversal of seen distances in comparison with the fixation-point: *i.e.* when first they were seen farther than the fixation-point, they appear then nearer and contrariwise.

In experiments with after-images, two vertical rods were placed in the median plane, one serving as a fixation-point and the other to be seen in double images. By fixating the rod for about 20–30 sec., *O* sees three images in the primary object and afterwards also in the after-images. When *O* turned his head and introduced a new fixation-point in the form of a rod with a pointed top, he saw three images of rods in complementary colors: the after-image of the fixation-rod at the distance of the new fixation and the after-images of double images nearer, farther, equidistantly, or changing similarly to the double images in the primary images. When the new fixation-rod is moved towards *O*, the after-images move with it, but in a retarded way, as if staying behind it for a while. (The impression is as if the after-images were dragged along by the fixation-rod.) At the same time, they diminish in size and come closer together laterally. When the new fixation-point is moved away from *O*, the opposite phenomena occur.

When after-images for double images have been produced, and one

[12] Helmholtz, *op. cit.*, 3, 1925, 400 ff.; Hering, Das Gesetz der identischen Seherichtung, *Du Bois-Raymonds Archiv*, 1864, 44 ff.; Anathon Aall, Über den Massstab beim Tiefensehen in Doppelbildern, *Z. Psychol.* 49, 1908, 108-127, 161-205; A. von Tschermak und P. Hoefer, Über binoculare Tiefenwahrnehmung auf Grund von Doppelbildern, *Pflügers Arch. ges. Physiol.*, 98, 1903, 299-321.

[13] K. N. Ogle, On the limits of stereoscopic vision, *J. exp. Psychol.*, 44, 1952, 253-259; Zając, op. cit., *Acta Psychol.* 12, 1956, 111-129.

eye is closed, then only what is seen by the other eye remains in the negative after-image, *i.e.* the monocular image of the fixated object and one double image, but the effect of changing size with changing distances remains.

After-images of double images are not something steady and unchanging as concerns their localization in the third dimension, they undergo changes similar to those of primary objects and of primary double images. With small number of experiments and with difficulties in measuring the corresponding times, the relations between the changes occurring in after-images and those in primary objects could not be ascertained. One must take into account the fact that the after-images do not come into being only when we turn our eyes from primary objects; they are already there after a few tenths of seconds of fixation, they develop while covering these primary objects, which thus change saturation of their color and distinctness. Their changes after being turned into pure after-images depend on the time of their duration, on the changes they undergo during their union with the primary objects, and on the nature and color of the background against which they are seen.

(*2 b*) The third group of investigations concerned the after-images of rods placed near the horopter passing through the actual fixation-rod. In that case they were seen by binocular vision as single; but when they were placed at greater geometrical distance from the horopter, several changes in the appearance of lateral rods occurred. They were more subject to rivalry and were sometimes seen as larger flat surfaces rather than as round rods; their observed lateral distance from the fixation-rod also changed. On the contrary, images of the lateral rods placed nearer to the horopter are steadier, *i.e.* less subject to rivalry effects.

When, for example, the central rod, which is fixated, is placed at 56.9 cm. from *O*, with the right rod at 62.3 cm., and the left one at 60.3 cm. (the lateral distance of the lateral rods from the fixation-rod is 5°), *O* (*SH*) reported, under binocular vision, that both lateral rods are seen singly and farther away than the fixation-rod, the right one farther than the left one. With prolonged fixation, he reported that the right image underwent more frequent changes of perceived distance than the left one. When the left rod was moved nearer by 2 cm. to 58.3 cm., the changes described above almost ceased to be observed in prolonged binocular fixation. It is easy to understand that the greater the distance from the horopter, the more the object is seen in overlapping double images although these are not distinctly discriminated by *O* as such; they are sometimes seen as flat, larger, sometimes as single, narrower images, changing their lateral

distance from the fixation-rod. These phenomena are frequently accompanied by changes in depth-perception—an alternating coming nearer and moving away. In all cases, after prolonged fixation the differences in distance between the fixation-point and images of lateral objects diminish and become less certain.

All these phenomena occur also in after-images, sometimes with greater distinctness and sometimes less pronounced. For example, one *O* (*SH*), in the situation described above with three rods (left one nearer than the right one, both farther than the fixation-rod), did not discriminate in after-images the difference in distance between the left and the right one, although he saw both farther than the fixation-rod.

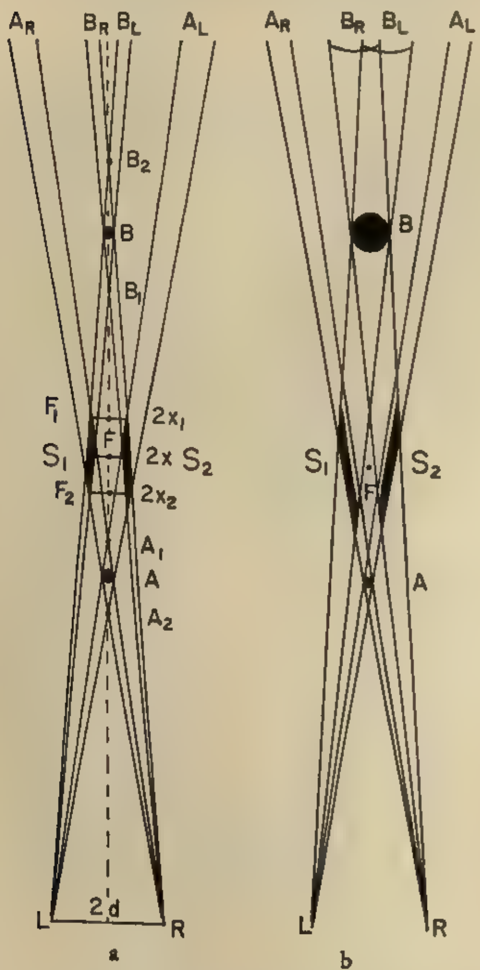
(*2 c*) In the fourth group of experiments the stereoscopic images in form of 'walls' served as primary objects. When we place three similar rods in the median plane and the middle one is moved to the position where the crossed images of the nearer rod overlap with uncrossed images of the farther rod, then instead of four images we see two stereoscopic images frequently of elongated form on either side of the fixation-point at the same distance with it and inclined to the median plane. When we move the fixation-point towards us, the stereoscopic images are seen farther away; when away from us, the 'walls' are seen nearer (see Figs. 2a and 2b). In several cases these images are seen as something like short 'walls,' 'ribbons,' or slanting surfaces, inclined to the median plane with their farther ends, when the monocular image of the nearer rod is larger than that of the farther one, and with their nearer ends, when the monocular image of the farther rod is seen as larger than that of the nearer one.[14] Not all of the *O*s see this relief of 'walls' at once, but their relative distance in comparison with fixation-rod is easily and distinctly observed.

When this kind of stereoscopic images served as primary objects for after-images, one easily observed, in the following after-images, the corresponding differences in distances between the after-image of the fixation-rod and the after-images of stereoscopic images. Also the relief of 'walls,' their elongated shape, and their inclination to the median plane were easily observed—if that was seen in the primary image.

A very interesting phenomenon of changed relief was observed by prolonged fixation in the case when the 'walls' were inclined to the median plane with their farther ends, especially when the nearer edges, were seen as equidistant with the fixation-point. Then inner parts of the 'walls' by

[14] For a description of 'wall-phenomena,' see Zając, Depth-perception of stereoscopic images resulting from fusion of crossed and uncrossed double images, this Journal, 72, 1959, 163-183.

prolonged fixation turned about the outer edges of the 'walls' and came in front of the fixation-point. This was observed also in other positions but not so easily as in that described above. After some time the 'walls' re-



FIGS. 2 a AND 2 b. RELIEF OF STEREOSCOPIC IMAGES IN FORM OF 'WALLS'

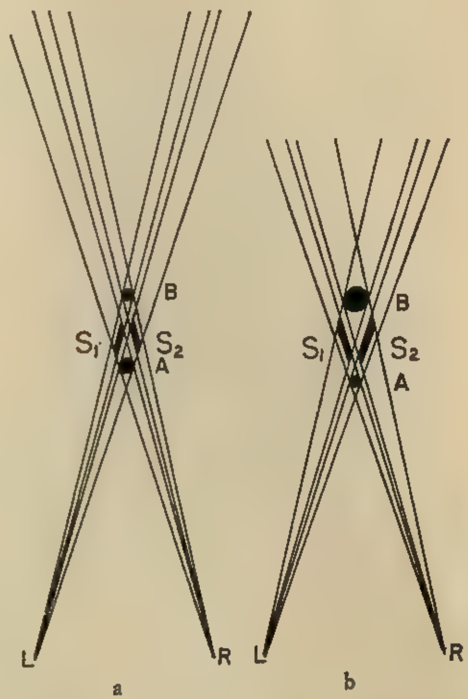
Fig. 2 a, fusion of images of the nearer rod seen as thicker with images of the farther rod seen as thinner; Fig. 2 b., fusion of images of the nearer rod seen as thinner with images of the farther rod seen as much thicker. 2 d = interocular distance; x = lateral distance from the median plane; other notations the same as for Fig. 1.

turned to their initial positions. The same phenomena of reverse relief were observed in after-images.

Within this group of experiments the possibility was explored of comparing the distance of after-images with the distance of primary objects. If, after a fixation of 20–30 sec. when stereoscopic images were used as

primary objects, *O* transfers fixation on a newly introduced rod on the right of the walls; then he easily observes the differences in distance of after-images with the primary objects. Thus we may ascertain whether it is possible to compare the perceived relative distances of after-images

Figs. 3 a and 3 b. "V"—SHAPED FORMS OF THE STEREOSCOPIC IMAGES
Fig. 3 a, inverted 'V'; Fig 3 b upright 'V' Notations the same as for Fig. 1.



with the real objects in the field of vision as well as with the images of stereoscopic character.

(*2 d*) Among others, I made experiments with a special kind of 'walls' assuming a normal or inverted shape of a 'V,' in which, between its two arms, a small interspace was seen (see Figs. 3 a and 3 b). This kind of 'wall' is produced with two rods placed in the median plane when, with monocular vision with either eye, both images of the rods are nearly touching each other. To produce the inverted 'V,' the distances used were 55.7 cm. for the nearer rod, 6 mm. in diameter, and 69.2 cm. for the farther rod of the same diameter; to produce the straight 'V,' the distances were 56.55 cm. for the nearer rod, 3 mm. in diameter, and 69.7 cm. for the farther rod, 6 mm. in diameter.

The relief of these 'walls' was very easily observed in primary images. Very rarely were inversions perceived. With prolonged fixation, however, flattening sometimes occurred. In after-images, the relief corresponding to what was seen in primary images, was also very easily observed and this on varying distances of the newly introduced fixation-point. When, however, images were thrown on a surface 1.5–2 m. distant, then the relief was lost and only flat, faded surfaces were perceived.

Some unexpected and interesting results were obtained, when a black or red rod was introduced in the lateral field of vision and seen in primary image at a distance of the middle of the 'V' or even at the far end of it: then the after-image of this laterally placed rod was seen distinctly as much brighter and distinctly nearer than the after image of 'V.' Only when this lateral rod was placed some distance behind the 'V,' it was seen in the after-image as equidistant or slightly behind the after-image of 'V.' This phenomenon was not investigated any further, but it can be of great importance to the theory of depth-perception of colors and thus deserves to be more carefully investigated. Perhaps it is an example of chromatic stereoscopy explained by Hartridge.[15] We have observed, however, that this change of depth-perception in after-images is the same for the dark as for the light background on which the after-images are seen. When the after-images of 'V' were transferred to the right or to the left of the original arrangement for 'V,' then similar observations were made as with other 'walls.'

Supplementary Observations

In the course of these investigations, the author and other *Os* observed movements of after-images when their positions in space were not secured by a steady fixation. It is known that after-images are very useful in investigations on eye-movements.[16] This is due to the fact that the retinal images corresponding to after-images are not displaced and that the after-images move only with the eyes.

When, after producing after-images, we stabilize their position by fixating a point on a wall for example and then transfer our regard to the space between the wall and the eyes by greater convergence we can then observe some characteristic movements of after-images different for different kinds of primary images and so discover various automatic or semi-automatic movements of the eyes. The direction of these movements is generally perpendicular to the extention of primary objects. For example, when vertical rods were used as primary objects, several *Os* reported pendular right and left movements of the after-images. The amplitude of these movements

[15] Hamilton Hartridge, *Recent Advances in the Physiology of Vision*, 1950; for Hartridge's investigation on color stereoscopy, see also Herbert Schober, *Das Sehen*, 2, 1954, 447-448.

[16] J. B. Listing, Beitrag zur physiologischen Optik, In Oswald's *Klasiker d. exp. Naturwissenschaft*, 147, 1905.; Helmholtz, *op cit.*, 3, 1925, 37 ff., 248 ff.

was about 10–15°, and their period was 1–2 sec. Sometimes one of the *O*s reported only a steady movement to the right or to the left. In certain conditions a saccadic movement to the right or from below upward was superimposed upon the pendulum movement.

When horizontal rods were used, the pendulum movement down and up was usually accompanied by upward saccadic movements. When, in this case, *O* moves his eyes voluntarily downward that the after-images are lowered, this lowering occurs in a steady, not a saccadic, movement; then the pendluum movement upwards recommences.

In the case of a circular object—an electric bulb for example—*O* (the author) rarely observed a pendular movement; he usually reported a steady movement upward and to the right.

With after-images produced from combined horizontal and vertical rods, the combined movements are generally perceived, but mostly the after-images move pendularly in the horizontal combined with saccadic, upward movements. When more attention is paid to the horizontal, the upwards movements predominate; when to the vertical, the pendular, horizontal movements prevail.

By a voluntary act, one *O* (the author), observed the after-images moving in ellipses to and away from him, as if he were playing with a yo-yo—but with this difference: the after-images diminished in size when approaching *O* and grew in size when going away from him, just opposite to what would be the case if he were playing with the yo-yo.

The number of observations was not large enough to allow for general conclusions in the matter of eye-movements; they indicate only that the automatic or semi-automatic movements of the eyes are sometimes determined by the shape of primary objects in the field of vision.

DISCUSSION

We do not know the origin of after-images. Are they of retinal or central origin, or conditioned by both? The important fact is that they are seen in the field of vision, in front of the *O*. They cannot be considered as 'objects,' as Emmert claims, but as phenomena, data, realities. Why they are seen, is and will probably long remain a mystery.

When we say they are 'projected' to the outside, as the projectionists' theories contend, we do not explain anything. It must be noted that the lines on which 'projection' would be effected are different from those by which the light-rays come into our eyes; the latter are refracted on every refracting surface outside and inside the eye, while the direction visual lines are understood to pass from the retinal image straight through the point on the last refraction-surface—or is it the second nodal point of the crystalline lens?—to the outside. The same is true about directions in which images of objects in the field of view are seen. To base a theory of vision on this evidence, one should first define what kind of process—physical or otherwise—is here involved.

When we use such words or notions as visual lines, visual axes, directional lines, visual angles, we must take into account the fact that they are abstractions or illustrations; we can use them only where they can define relations between what we perceive and geometrical gauges of objects or of other realities outside. Where these relations can be illustrated by lines, angles, or graphs, the latter have their use and value, but they certainly do not explain any phenomena of spatial localization.

The most interesting aid here, beside directional lines, is the 'notion' of visual angle as defined above (p. 509).[17] Images of objects in the lateral field of view, double images of stationary objects grow when the fixation-point is moved away from *O* and diminish when the fixation-point is brought nearer. This happens in monocular as well as in binocular vision. In certain cases these changes in size are accompanied by changes in perceived distance of these objects or images. This law of visual angle is the same as the so-called Emmert's law for after-images, with certain qualifications. The investigations have proved that Emmert's law is accurate only within certain limits, and for certain shapes of after-images. The discrepencies in sizes of after-images of objects of equal geometrical height (Hindenburg's head and a vertical line in Noll's experiments for example)[18] might in author's opinion be attributed to the differences in accommodation when different shapes of the objects are involved and thus differences in size of retinal images occur. For example, the state of accommodation for a vertical line might be different than the accommodation for Hindenburg's head of the same height. The result might be that the retinal image of the vertical line would be higher than that of Hindenburg's head and thus also the after-image of the vertical line would be higher than the after-image of the head. In this way, after-images could perhaps be helpful in discovering and measuring the changes of accommodation for different images of objects. Thus the size of after-images is determined not by the size of an object but by the size of retinal image and by the distance of the fixation point, and, with that qualification, the law of visual angle could be applied to after-images. It must be noted however, that the perceived size of after-images is quite a different matter. Emmert's law and the law of visual angle are concerned with measured geometrical sizes which grow proportionally to the distance of the fixation-point from *O*; the perceived changes cannot be measured, they can only be compared with the sizes of images of other objects in the field of view; the fact of grow-

[17] The arms of visual angles are formed by directional lines, which denote directions in which the edges or contours of objects are seen.

[18] Adolf Noll, Versuche über Nachbilder, *Psychol. Forsch.*, 8, 1926, 3-27.

ing or diminishing size can be perceived with moving fixation-point but the amount of the change can only be estimated. That is the case of after-images, but it is not the same with real objects in the field of vision. When we change the position of the fixation-point within the range of accommodation of the eyes, another factor has to be taken into account, *i.e.* accommodation. The greater the distance of the fixation-point, the less curved or more flattened becomes the crystalline lens, the greater are the retinal images of the stationary objects placed in the lateral field of view: they are seen as greater.[19]

This fact is often not taken into account for explanation of different changes in perceived size of stationary objects with changes of distances of fixation-point. It seems that some authors assume that the retinal images of those laterally situated objects do not change at all and try to explain the perceived changes in another way (by constancy for example).[20]

Another kind of role of accommodation is to be noted here. When the objects in the lateral field of view come outside the so-called 'depth of focus', their retinal images become blurred and thus generally larger, than the sharp, distinct ones. In normal vision all three factors—visual angle, curvature of the lens, and blurredness or distinctness of the retinal image—play their role in determining the size of the images; in after-images the visual angle only operates. This difference can be illustrated by using a magnifying glass. Look, for example, at a primary object (a small black square on a light yellow wall) for 20–30 sec. When a negative after-image is formed, look at the primary object through a magnifying glass. You will see the original square magnified and beside it, or overlapping upon it, the negative after-image in its unchanged size. The magnifying glass has no influence on the size or the position of the after-image, but it changes the size of a real object in the field of vision.

Then can we say that accommodation does not play any role in perceiving the size and distance or relief of after-images? The accommodation of the eye is determined, as we know, when the fixation-point is chosen in the field of view. This is so within the range of accommodation. The after-images are seen at a distance of the fixation-point and their size depends on the distance of this fixation-point. This means that a relation exists between the size of the after-image and its relative distance and the

---

[19] Zając, *La localization en profondeur des image doubles*, 1923, 97 ff.; op. cit., *Brit. J. Psychol.*, 44, 1953, 132-144.

[20] A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance variant, this JOURNAL, 54, 1941, 21-37; J. J. Gibson, *Perception of the Visual World*, 1950, 23 ff., 163 ff., 175.

actual state of accommodation of the eye; but it does not mean that the distinctness or blurredness of the after-image can be changed by a change of accommodation. When, for example, the new fixation-point is chosen outside the range of accommodation, it is seen blurred, but not so the after-image, which was produced by looking at distinctly perceived objects. With growing distance of the fixation-point beyond certain limits, different for eyes of different *Os*, the after-images growing in size become so to speak diluted and are seen rather blurred, but that cannot be considered as due to the influence of the actual state of accommodation. This is what we can say about relation between after-images and accommodation in monocular vision.

What about the role of convergence with binocular vision? The binocular after-images of objects, which were fixated, cannot be seen in the lateral field of view, and the after-images of objects seen in the lateral field of vision cannot be seen in the center of the field of view. One cannot dissociate an after-image of fixated objects into two double images, or fuse after-images of double images, by any artificial means in the form of a pressure on the eyeball. By pressure on the eye one can change the color or make the after-images disappear, but one cannot change its size or a lateral distance between the after-images of double images or dissociate a single after-image into two images. The perception of the distance might be lost due to the fact that every object in the real field of view is then doubled, and no one image can become a binocular fixation-point.

These facts might be considered as supporting the theory of corresponding points of retinas and of the role of binocular disparity, as well as of the law of intersection of directional visual lines, stating that geometrical distance of stereoscopic images as defined by that intersection. In any case after-images cannot be used against the theory of corresponding points on retinas or against the projection theory of vision, as has been done, without foundation, by Hering.[21] The point of convergence on the fixation-point is the point where we see after-images of the fixation-point of the primary object and the geometrical distance of the fixation-point determines also the size of the whole after-image. The objects near the horopter and with a small binocular disparity are seen in single after-images with difference of distance, generally corresponding to what was seen in primary objects. Also, if real objects or stereoscopic images are seen in relief, the after-images within certain limits of distance are also seen in relief. The changes occurring in reliefs and in perceived distances with prolonged

[21] Hering, *Beiträge zur Physiology*, 2, 1862, 132 ff.

fixation are also reproduced in after-images. One cannot, as yet, establish a fixed relation between changes perceived in primary objects and changes in after-images. One can only state that similar changes occur in after-images as in primary objects or images. This problem is important because findings concerning relations between these two phenomena could give better insight into the processes occurring on the retinas and in the whole organ of vision.

Summary

The present investigation was conducted to establish facts concerning localization of after-images in the third dimension. The first conclusion based upon our results is that after-images of fixation-points in primary objects are seen at the distance of the new point of fixation, and that the size of after-images varies with the distance of this point. When the fixation-point is moved, the after-images move with it. These results stand in contradiction to Hering's contention that the size of the images varied independently with the distance of the fixation-point. Our results also established that Emmert's law must be interpreted in such a way, that *the sizes of after-images are determined not by the distance of the background on which they are seen, but by the distance of the fixation-point.*

Several cases were investigated, using as primary objects real objects of colored rods, double images, and stereoscopic images produced by natural (free) stereoscopy. Generally similar phenomena and changes were observed in after-images as concerns depth-perception as in primary objects; this was observed with monocular as well as with binocular vision, but relations between what occurs in primary objects and in after-images could not be established. On several occasions, however, when colored objects were used, one observed different reliefs in after-images than in primary objects.

Observations made in these investigations have many implications concerning general theory of depth-perception, as well as the role of convergence, accommodation, visual angle, and other factors, which are considered in the discussion.

In the course of these investigations, some observations on movements of after-images and of the eyes were made when their position in space was not held by steady fixation. When the position of after-images is first stabilized by fixation upon a point on a wall, and then our regard is transferred between the wall and the eyes by a change in convergence of the eyes, we observe some characteristic movements which are different for different kinds of images. The direction of these movements is gen-

erally perpendicular to the extention of primary objects. With the after-images produced from combined horizontal and vertical rods, the combined movements are generally observed: after-images move chiefly in the horizontal direction, like a pendulum combined with upward saccadic movements. When attention is concentrated on the horizontal parts, the upward movements predominate; when concentrated on the vertical, the pendular, sidewise movements predominate. The phenomena indicate that the automatic or semi-automatic movements of the eyes are determined at times by the shape of primary objects.

THE EFFECT OF MUSCULAR INVOLVEMENT ON SENSITIVITY: ASYMMETRICAL CONVERGENCE ON THE DISTRIBUTION OF VISUAL SENSITIVITY

By Joseph H. McFarland, Heinz Werner, and Seymour Wapner, Clark University

The different distributions of sensitivity in the visual field for scotopic and photopic vision have been known for some time. For photopic vision, there is a high of sensitivity in the center of the visual field and decreasing sensitivity towards the periphery. With scotopic vision, however, the distribution of sensitivity is radically altered: there is blindness in the center of the visual field, increasing sensitivity up to 10° of visual angle right and left of the center and then decreasing sensitivity toward the periphery.[1] Since Helmholtz, the explanation of the distribution of sensitivity in the visual sensitivity does not always correspond to th edistribution of visual cones in the retina. If there is a high sensitivity at one point in the visual field, then this point projects to a retinal area high in receptor density.[2]

For some time, however, it has been known that the distribution of visual sensitivity does not always correspond to the distribution of visual receptors in the retina. Both temporary and relatively permanent changes in this correspondence between receptors and sensitivity have been reported.

Temporary changes. Lipp has reported, for example, that if *S*s fixate one point in the visual field while attending to another point, the point attended to is more clear.[3] Dallenbach and others have shown that the location of the highest point of clarity in the visual field depends on such stimulus-properties as intensity, form and size.[4] These findings indicate that the highest point of sensitivity is not always the fovea which has the highest density of cones.

---

* Received for publication August 18, 1959. This investigation was supported by research grant M-348 from the National Institute of Mental Health, Public Health Service.

[1] Th. Wertheim, Über die indirekte Sehschärfe, *Z. f. Psychol.*, 7, 1894, 172-187.

[2] Hermann von Helmholtz, *Treatise on Physiological Optics* (J. P. C. Southall, ed.), ed., 2, 1924, 386; C. H. Graham, and R. Granit, Comparative studies on the peripheral and central retina, *Amer. J. Physiol.*, 98, 1931, 664-671; Selig Hecht, C. Haig and G. Walk, The dark adaptation of retinal fields of different size and location, *J. gen. Physiol.*, 19, 1935, 321-337; J. L. Falk, Theories of visual acuity and their physiological bases, *Psychol. Bull.*, 53, 1956, 109-133.

[3] O. Lipp, Über die Unterschiedsempfindlichkeit im Sehfelde unter dem Einflusse der Aufmerksamkeit, *Arch. f. d. ges Psychol.*, 19, 1910, 313-394.

[4] J. N. Curtis, and W. S. Foster, Size vs. intensity as a determinant of attention, this Journal, 28, 1917, 293-296; L. G. Meads, Form vs. intensity as a determinant of attention, this Journal, 26, 1915, 150-151. K. M. Dallenbach, Position vs. intensity as a determinant of clearness, this Journal, 34, 1923, 282-286.

Permanent changes: (a) Normal. One of Dallenbach's studies indicates that right-handed subjects see a point to the left of fixation as most clear.[5]

(b) Abnormal. Many of the strabismic patients examined by Worth had developed what he called a 'pseudo-fovea.'[6] For example, if the right eye turned outward in fixation, the point of greatest clarity for the right eye was located temporally from the anatomical fovea. Fuchs' study of hemianopic patients indicates that in certain cases of homonymous hemianopsia, a 'pseudo-fovea' develops also.[7] In these cases, the highest point of clarity is shifted away from the non-functional area and hence away from the anatomical fovea.

Various interpretations, other than strictly neuro-anatomical, have been proposed in the attempt to explain these facts. Lipp and Titchener interpret temporary changes in location of the high point of clarity in terms of central changes in the direction of attention.[8] Worth,[9] as does Linksz, relates the development of a 'pseudo-fovea' in the strabismic eye to a "defect of the fusion faculty."[10] Von Tschermak-Seysenegg interprets the 'pseudo-fovea' as the result of a cooperation of sensory and muscular anomalies, *i.e.* a changed location of the high point of sensitivity and a change in the direction of straight-ahead gaze.[11]

Goldstein attempts to handle both temporary and permanent changes. For Goldstein, the temporary changes in location of the high point of clarity are interpreted in terms of a dynamic process which includes the pattern of stimulation on the retina as well as the general attitude of the organism. Thus, either changes in the stimulus or the attitude of the organism may have consequences for the location of the high point of clarity. With respect to permanent changes, Goldstein assumes that "optimal visual performance" is based on a visual field organized around a center. Thus, the development of a 'pseudo-fovea' in hemianopic patients is interpreted as contributing to the "maintenance of optimal performance": in these patients, with blindness in half of the visual field, a 'pseudo-fovea' develops some distance from the blind area; this allows a spatial organization of sensitivity with a high of clarity at one point and decreasing clarity *right and left* of this point, which is the characteristic spatial distribution of sensitivity for normal vision.[12]

Problem. The aim of this paper is to consider an organismic factor not customarily mentioned in discussions of sensitivity; namely, muscular involvement. Specifically, we wish to consider the role of muscular changes in the ocular system in relation to changes in the distribution of visual sensitivity.

---

[5] Dallenbach, Attributive vs. cognitive clearness, *J. exp. Psychol.*, 3, 1920, 183-230.

[6] C. A. Worth, *Squint: Its Causes, Pathology and Treatment*, 1906, 165-166.

[7] W. Fuchs, Selections 28 and 30, in Willis D. Ellis (ed.), *A Source Book of Gestalt Psychology*, 1938, 333-343, 357-361.

[8] Lipp, *op cit.*, 313-394; E. B. Titchener, *An Outline of Psychology*, 1897, 125-135.

[9] Worth, *op cit.*, 165-166.

[10] Linksz, *Physiology of the Eye*, 2, 1952, 552.

[11] Arthur von Tschermak-Seysenegg, *Introduction to Physiological Optics*, 1952, 196.

[12] Kurt Goldstein, *The Organism*, 1939, 53-55.

In everyday situations, we tend to look directly at objects of interest. This may be accomplished by body, head, and eye movements, either singly or in combination. Regardless of the manner, this activity of looking directly at objects of interest has three consequences for perception: (1) the object is localized in space; (2) the object is localized with reference to the viewer (egocentric localization); and (3) the object is seen with maximal clarity, *i.e.* is seen more clearly than other objects present in the field of vision. Ordinarily when we change our direction of view, this change does not alter the location of the objective straight-ahead (objective median plane) or of objects localized with respect to the straight-ahead. What does change, when we look to one side of the objective straight-ahead (objective median plane), is the spatial distribution of clarity. When we look at objects straight-ahead, clarity is high for these objects and is low for objects right and left. When, however, we look at an object outside of the straight-ahead (objective median plane), clarity is now high for this object and low for other objects right and left of the object now viewed. This is in accord with the expectation that when one looks directly at an object, this object is centered on the fovea, and hence should be seen with maximal clarity.

It should be emphasized that, under certain conditions, the point of highest sensitivity is *not* located at the fovea: an object stimulating an extra-foveal location may be seen as most clear. For example, in the Lipp experiment reported previously, *S*s looked directly at a point objectively straight-ahead, but were instructed to attend to an object which stimulated an extra-foveal location. These *S*s reported that the object *attended to* appeared clearer than other objects visible in the field and, further, that *this object was located straight-ahead.* For these *S*s, an object stimulating an extra-foveal location on the retina is seen *as if* it were objectively straight-ahead.[13] A further instance of this is the development of a 'pseudo-fovea' in strabismic and hemianopic patients. For these patients, respectively, one or both of the eyes are slightly turned, such that objects located objectively straight-ahead, *i.e.* in the objective median plane, stimulate extra-foveal locations on the retina. These *S*s report that objects objectively straight-ahead, although stimulating extra-foveal locations, appear clearer, than other objects visible and, further, appear straight-ahead.[14] For these *S*s then, objects in space stimulating an extra-foveal location in the retina are seen in a manner characteristic of objects which stimulate the fovea during straight-ahead gaze in normals.

[13] Lipp, *op. cit.*, 313-394.
[14] Worth, *op. cit.*, 165-166; Fuchs, *op. cit.*, 333-343, 357-361.

In addition to showing that the highest point of clarity does not always coincide with the fovea, these findings show a concurrent reorganization of visual space, *i.e.* maximal clarity is displaced right or left of the fovea and the apparent straight-ahead (apparent median plane) is also displaced right or left of the fovea. This suggests that in addition to the anatomical factors discussed earlier, a comprehensive approach to the problem of the distribution of visual sensitivity must consider the conditions which organize perceptual space. Evidence has been accumulating that muscular involvement reorganizes perceptual space, *i.e.* displaces the apparent median plane.[15] Thus we might expect that muscular involvement would also reorganize the distribution of visual sensitivity. Specifically, it is expected that increased convergence of the right or the left eye will change the distribution of visual sensitivity. Further, it is hypothesized that increased convergence of the left eye (which displaces the apparent median plane to the left of the objective median plane) will increase clarity relatively to the left of the objective median plane.[16] Conversely, increased convergence of the right eye (which displaces the apparent median plane to the right of the objective median plane) is expected to increase clarity relatively to the right of the objective median plane.[17]

Two experiments were designed to test these hypotheses. These experiments differed in one major respect: in the first, visual stimuli were presented under scotopic illumination; and in the second, they were presented under photopic illumination.

Experiment I. Scotopic Conditions

Subjects. Sixteen *Ss* were tested within an 8 × 8 replicated Latin Square design. All were students.

Apparatus and procedure. S was dark-adapted for 10 min. and the stimulus-patterns were then exposed to him in a dark room. Four patterns were used. Each consisted of a *pair* of luminous Landolt rings with a pin-point of orange light centered between them which he was instructed to fixate. Head-position was kept in the objective median plane by means of a head-rest and each stimulus-pattern was so presented that the fixation-point was located in the objective median plane, viz. stimulating the fovea. There was a break in each of the two rings constituting the stimulus-pattern. Breaks in the rings were either 24′ or 1° 4′ of visual angle left and right of fixation. Each ring is 40′ of visual angle in diameter (see Fig. 1). The stimulus-patterns were presented at a distance of 11 ft. A threshold of clarity was determined in the following manner. Each stimulus-pattern was presented for 0.04 sec. at increasing intensities of illumination. This continued until the break in both

[15] J. H. Bruell, *Visual Egocentric Localization; An experimental study,* unpublished doctoral dissertation, Clark University, 1953.

[16] Bruell, *op. cit.*, 52-60.

[17] Bruell, *op cit.*, 52-60.

left and right rings was seen in its correct location. The stimulus-patterns were presented under two conditions to all *Ss*; once while wearing a 10°, base-out prism in front of the left eye, which induces greater convergence of the left eye; and a second time, while wearing a 10°, base-out prism in front of the right eye, which induces greater convergence of the right eye. The prism displaces the orange pin-point of light, which is fixated, 10° temporally on the retina of the eye wearing the prism. For the displaced pin-point of light to be fused with the pin-point of light stimulating the other eye, the eye viewing through the prism must turn nasally:

FIG. 1. PAIRS OF LANDOLT RINGS USED AS STIMULUS-PATTERNS IN EXPERIMENTS I AND II

under these conditions the pin-point of light is under binocular fixation, that is, it now stimulates both foveas. After this binocular fusion has occurred, the eye viewing through the prism is under greater convergence than the eye without the prism.

Using this method, we obtained thresholds for the four stimulus-patterns, in terms of the illumination intensity of the patterns required for correct localization of the break in both left and right Landolt rings.[18]

Results. The data were analyzed using difference-scores. Arbitrarily, a minus (−) difference indicates greater clarity for the left ring and a plus (+) difference greater clarity for the right ring. An analysis of variance, utilizing these scores, and the means for the experimental conditions are presented in Table I. Examination of this table shows that the prisms yielded significant sources of variance. Inspection of the means shows that when the base-out prism is in front of the left eye, clarity is relatively greater for the left ring (left of the objective median plane) than when the base-out prism is in front of the right eye. When the base-out prism is in front of the right eye, clarity is relatively greater for the right ring (right of the objective median plane) than when the base-out prism is in front of the left eye. Table II presents the thresholds for each stimulus-pattern for both positions of the prism. The means are in variac readings. It can be seen

[18] Selig Hecht, The relation between visual acuity and illumination, *J. gen. Physiol.*, 11, 1928, 225-281.

from Table II that the left-right relative clarity of each pattern changes systematically and in the same direction with each of the experimental conditions: when the prism is in front of the left eye, there is *relative* greater

TABLE I

ANALYSIS OF VARIANCE UTILIZING DIFFERENCES IN ILLUMINATION BETWEEN LEFT AND RIGHT RINGS AT WHICH CORRECT LOCALIZATION OF THE BREAK OCCURRED

(Experiment I)

Source of variation	*df.*	Sum of squares	Mean squares	*F*
Individuals (I)	15	2,055	137.00	<1.00
Sequence (Seq)	7	499	71.28	<1.00
I×Seq	8	1,556	194.50	—
Order	7	3,387	483.86	1.52
Stimulus Patterns×Prisms	7	10,193	1,456.14	4.58*
Prisms (P)	1	2,032	2,032.00	6.39†
Stimulus Patterns (S)	3	8,007	2,669.00	8.40†
P×S	3	154	51.33	<1.00
Square Uniqueness	42	9,431	224.57	<1.00
Error	56	17,801	317.88	—
Total	127	42,322	—	—

Base-out prism in front of left eye, −5.47; Base-out prism in front of right eye, +2.50‡.

* Significant at the 1% level. † Significant at the 5% level.

‡ Plus (+) indicates a lower threshold for the right ring, minus (−) indicates a lower threshold for the left ring.

TABLE II

MEAN INTENSITY OF ILLUMINATION AT WHICH CORRECT LOCALIZATION OF THE BREAKS FOR EACH STIMULUS-PATTERN OCCURRED

(Magnitudes are variac readings, a low magnitude indicates a low threshold)

(Experiment I)

Stimulus-pattern	Prism	Left ring	Right ring	L–R*
1	Left eye	71.56	87.18	−15.62
	Right eye	75.31	83.44	− 8.13
2	Left eye	71.56	73.12	− 1.56
	Right eye	79.69	73.44	+ 6.25
3	Left eye	77.81	86.25	− 8.44
	Right eye	87.81	90.62	− 2.81
4	Left eye	73.12	70.00	+ 3.12
	Right eye	80.31	65.62	+14.69

* Plus (+) indicates a lower threshold for the right ring, minus (−) indicates a lower threshold for the left ring.

clarity for the left ring (left of the objective median plane) than when the prism is in front of the right eye. Conversely, when the prism is in front of the right eye, there is *relative* greater clarity for the right ring (right of

the objective median plane) than when the prism is in front of the left eye. This is true for each of the stimulus patterns.

Fig. 2 presents the means of the different locations tested on the horizontal meridian of the eye. It can be seen that the distribution of clarity for both experimental conditions meets expectations concerning scotopic

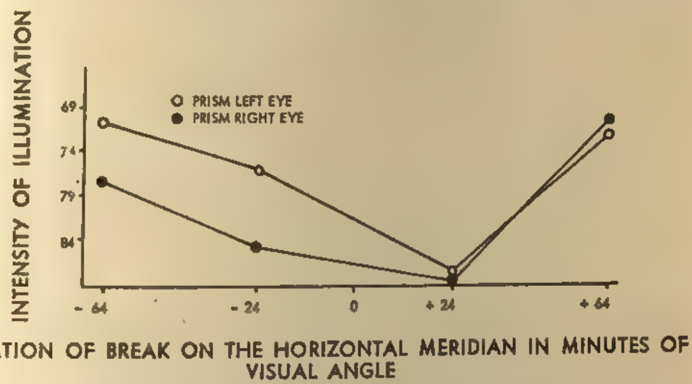


FIG. 2. EXPERIMENT I: MEAN THRESHOLDS OF SENSITIVITY FOR THE FOUR LOCATIONS UNDER SCOTOPIC CONDITIONS
('Minus' is left of fixation and 'plus' is right.)

conditions of illumination: there is a low of clarity in the center and increasing clarity right and left of the center.[19]

This figure also shows that the major effect of the experimental conditions is in the left half of the visual field. The effect of the experimental conditions is manifest, however, in both left and right halves of the visual field tested. When the prism is in front of the left eye, clarity is greater in the left half of the visual field; and when the prism is in front of the right eye, clarity is higher in the right half of the visual field. The one exception to this is at 24′ of visual angle right.

EXPERIMENT II. PHOTOPIC CONDITIONS

Subjects. Sixteen *S*s were tested within an 8 × 8 replicated Latin Square design. All were students.

Apparatus and procedure. This experiment repeated the first in all respects except the following: *S* was not dark-adapted; the experimental room was dimly lit; and the various stimulus-patterns (Fig. 1) were presented on an illuminated gray screen. In addition, fixation was strictly controlled in the following manner. *S* was instructed to fixate a black dot on the illuminated screen while looking through

[19] Wertheim, *op cit.*, 172-240.

binocular apertures. These apertures were approximately 5° of visual angle in diameter. In front of these apertures, two independently adjustable panes of glass were mounted. On each pane a small black dot was fixed. One of these panes was so adjusted that the dot coincided with the line of regard of the right eye; and the other pane, that the dot coincided with the line of regard of the left eye. *S* would see three near dots when fixation was maintained on the far fixation-dot without the binocular apertures. When *S* viewed the fixation-dot through the binocular apertures, the right and left dots were occluded from his vision by the apertures. By this means, any change in fixation would result in *S*'s seeing more than one dot.

Before each presentation of a stimulus-pattern, *S* was asked to report the number of dots seen. No exposure was made unless only one dot was seen. Further, if *S* saw more than one dot while a stimulus-pattern was presented, the pattern was presented again at the same intensity of illumination. In this way all stimulus-patterns were presented in the same retinal location for both experimental conditions.

TABLE III

Analysis of Variance Utilizing Differences in Illumination Between Left and Right Rings at Which Correct Localization of the Break Occurred

(Experiment II)

Source of Variation	*df.*	Sum of squares	Mean squares	*F*
Individuals (I)	15	7,310.78	487.39	<1.00
Sequence (Seq)	7	2,765.47	395.07	<1.00
I×Seq	8	4,545.31	568.16	—
Order	7	615.47	87.92	<1.00
Stimulus Patterns×Prisms	7	1,577.97	225.42	<1.00
Prisms (P)	1	361.18	361.18	1.53
Stimulus Patterns (S)	3	269.38	89.79	<1.00
P×S	3	947.41	315.80	1.34
Square Uniqueness	42	11,423.75	271.99	1.15
Error	56	13,142.19	234.64	—
Total	127	34,070.16	—	—

Base-out prism in front of left eye, −1.48; base-out prism in front of right eye, +1.88.*

* Plus (+) indicates a lower threshold for the right ring, minus (−) indicates a lower threshold for the left ring.

There was considerable variation between individuals in their ability to perform this task. Some of the *Ss*, while wearing the base-out prism never reported diplopia either before or during presentation of the stimulus-patterns. Further, many of the *Ss* required training in this task. For this purpose, *S* fixated and followed *E*'s finger while it was moved near and far in the field of vision. Some of these *Ss*, however, were still unable to perform the task after 20 min. of training, hence were dropped from the study.

Results. The data were analyzed in the same manner as in Experiment I. An analysis of variance and the means for the experimental conditions are presented in Table III. Examination of this table shows that prisms are not significant sources of variance. Inspection of the means show, however, that when the base-out prism is in front of the left eye, clarity is *relatively*

greater for the left ring (left of the objective median plane) than when the base-out prism is in front of the right eye. When the base-out prism is in front of the right eye, clarity is *relatively* greater for the right ring (right of the objective median plane). The direction of these mean differences is the same as for Experiment I. Table IV presents the thresholds for each stimulus-pattern for both prism conditions. The means are in variac readings. It can be seen from Table IV that for Stimulus-Patterns 1, 2, and 3, when the prism is in front of the left eye, there is *relative* greater clarity for the left ring (left of the objective median plan) than when the prism

TABLE IV

Mean Intensity of Illumination at Which Correct Localization of the Breaks for Each Stimulus-Pattern Occurred

(Magnitudes are variac readings, a low magnitude indicates a low threshold)

(Experiment I)

Stimulus-pattern	Prism	Left ring	Right ring	L–R*
1	Left eye	91.25	97.19	−5.94
	Right eye	93.44	89.06	+4.38
2	Left eye	90.31	93.12	−2.81
	Right eye	89.06	87.50	+1.56
3	Left eye	92.81	90.94	+1.87
	Right eye	95.00	90.31	+4.69
4	Left eye	90.62	89.06	+1.56
	Right eye	93.12	95.62	−2.50

* Plus (+) indicates a lower threshold for the right ring; minus (−) indicates a lower threshold for the left ring.

is in front of the right eye. Conversely, when the prism is in front of the right eye, there is *relative* greater clarity for the right ring (right of the objective median plane) than when the prism is in front of the left eye. The effect of the prism conditions for these stimulus-patterns is in the same direction as in Experiment I. For Stimulus Pattern 4, however, the effect of the prism conditions is reversed.

Fig. 3 presents the means of the different locations tested on the horizontal meridian of the eye. It can be seen from this figure that the distribution of clarity for both prisms meets expectations concerning photopic conditions of illumination.[20] There is greater clarity toward the center and decreasing clarity right and left of this high point. Fig. 3 shows, however, that the distribution of clarity is drastically changed for the two prisms. With the exception of 1° 4′ of visual angle right, the effect of the prisms

[20] Wertheim, *op. cit.*, 172-240.

is manifest in both left and right halves of the visual field tested: when the base-out prism is in front of the left eye, clarity is greater in the left half of the visual field; and when the base-out prism is in front of the right eye, clarity is greater in the right half of the visual field.

Discussion

These experiments tested two hypotheses: (1) muscular involvement, that is, increased convergence of the left or right eye changes the distribu-

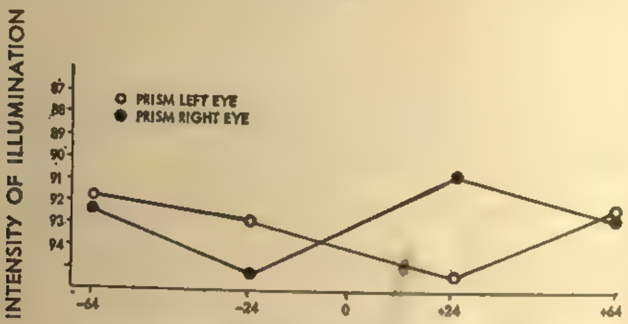


Fig. 3. Experiment II: Mean Thresholds of Sensitivity for the Four Locations Under Photopic Conditions
('Minus' is left of fixation and 'plus' is right.)

tion of visual sensitivity; (2) increased convergence of the left eye increases visual clarity *relatively* to the left of the objective median plane and conversely, increased convergence of the right eye increases visual clarity *relatively* to the right of the objective median plane.

In Experiment I, with scotopic illumination and fixation, it was found that increased convergence of the left or right eyes significantly changes the distribution of visual sensitivity. Further, it was found that there was a directional difference in the change depending on whether there was increased convergence of the left or the right eye: increased convergence of the left eye increased the visual sensitivity of the left half, relative to the right half, of the visual field, *i.e.* left of the objective median plane. Conversely, increased convergence of the right eye increased the visual sensitivity of the right half, relative to the left half, of the visual field, *i.e.* right of the objective median plane.

Experiment II differed from Experiment I in the use of photopic illumi-

nation and strict fixation. In this experiment, for three out of the four stimulus-patterns, visual sensitivtiy changes with muscular involvement in the same direction as in Experiment I. Changes in visual sensitivity for one of the stimulus-patterns were in an opposite direction.

The results of these experiments provide support for both of our hypotheses: (1) Asymmetrical muscular involvement of the ocular system changes the distribution of visual sensitivity for both rod and cone vision. (2) The direction of change for a particular location in the visual field depends, in general, on whether the left or the right eye has increased convergence.

These conclusions can be made because one can assume that the point fixated by the *S*s stimulates both foveas and thus the stimulus patterns are presented in identical retinal locations. Under the conditions of these two experiments, an anatomical explanation of clarity in terms of receptor density would predict no difference between the experimental conditions. Thus, our results require an interpretation which takes into account factors other than purely peripheral, *i.e.* retinal.[21] Presently, only the theory of Marshall and Talbot considers more central factors, that is, lateral and vertical summation, temporal and spatial dispersion of excitation in the striate cortex, and dynamic summation of excitation in the striate cortex due to nystagmus. Accordingly, a high peak of excitation in the striate cortex, provided either by a great number of receptors or vertical summation is the functional basis for foveal-peripheral differences in sensitivity.[22]

Though this view of sensitivity considers factors more central than the rods and cones, it remains within classical theory in so far as these factors relate to the sensory aspect of the ocular system. Our results indicate that in addition to the classical sensory functioning of the ocular system, the role of proprioceptive stimulation involving the motor aspect of the ocular system must be considered in a comprehensive theory of visual sensitivity.

An adequate interpretation of our results has to assume that motor and sensory functioning have a common field of interaction. This means that proprioceptive and exteroceptive stimulations, though coming from peripherally distinct sources, must functionally converge at some more central level of the organism. Granting, then, that proprioceptive and exterocep-

---

[21] The authors do not wish to give the impression of denying the contribution of retinal anatomical factors. The contribution of peripheral conditions such as the distribution of receptors on the retina to the total effect described in this paper is a complex problem which needs further study.

[22] W. H. Marshall and S. A. Talbot, Recent evidence for neutral mechanisms in vision leading to a general theory of sensory acuity, in H. Klüver (ed.) *Visual Mechanisms*, 1942, 117-164.

tive stimulations must converge at some central field, one may further ask whether one type of stimulation acts upon the other or is there true mutual interaction. The experiments reported indicate only that proprioceptive stimulation acts on exteroceptive stimulation. Other experiments, however, have shown that exteroceptive stimulation also effects proprioceptive stimulation through changing body tonus. For example, Chandler has reported that homogeneous light stimulation of either right or left eyes results in either left or right head turning.[23] On this basis, it can be assumed that there is a true mutual interaction between proprioceptive and exteroceptive stimulation.

This discussion explicates two requirements that must be met by an explanation of our results. First, the explanation must assume that motor and sensory systems are functionally integrated at some level of the nervous system. Secondly, the explanation, in its more general statement, must assume that motor and sensory stimulation may interact.

Previously, reference was made to the evidence showing a concurrent reorganization of visual space. Our results are in accord with these findings. With increased convergence of the left eye, which is known to displace the apparent straight-ahead to the left of the objective straight-ahead, visual sensitivity is relatively greater to the left of the objective straight-ahead. Conversely, with increased convergence of the right eye, which is known to displace the apparent straight-ahead to the right of the objective straight-ahead, visual sensitivity is relatively greater to the right of the objective straight-ahead. The findings indicate that both the distribution of visual sensitivity and the organization of visual space are dependent on muscular involvement, specifically asymmetrical muscular patterning.

Summary

Two experiments have been reported which show effects of asymmetrical convergence on the distribution of visual sensitivity for scotopic and photopic conditions of illumination. The results of these experiments also indictate that the change in the distribution of visual sensitivity is paralleled by a reorganization of visual space, viz. displacement of the apparent median plane. Two requirements for an interpretation, other than one that is strictly retinal, were explicated: (1) sensory and motor systems are functionally integrated at some level of the nervous system; and (2) there is a true mutual interaction between motor and sensory systems.

[23] K. A. Chandler, *The Effect of Moving and Non-moving Visual Stimuli Upon Head Torsion*, unpublished doctoral dissertation, Clark University, 1953.

GEOMETRICAL FACTORS IN ILLUSIONS OF DIRECTION

By Magloire Maheux, John C. Townsend, and Clement J. Gresock, Catholic University of America

Woodworth suggests that geometrical-optical illusions appear to be explained by pregnance.[1] The present authors believe that it is also necessary, in the case of *illusions of direction*, to take into account two more particularly fundamental factors; namely, *displacement* and the *Gestalt principle of direction.*

When one looks at Zöllner's or Wundt's figures (Fig. 1), parallel lines seem no longer parallel but *displaced.* In addition, each segment (Fig. 1, a,b,c,d,e,f,g,h), determined on the parallels by the cutting lines of the field, appears to deviate,

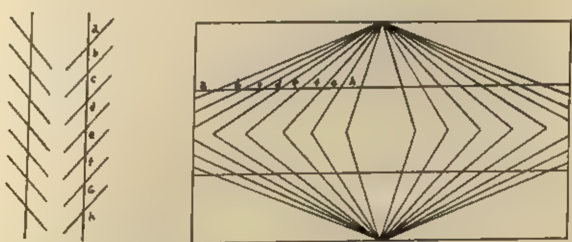


Fig. 1. Simplified Zöllner and Wundt Illusion

i.e. to turn around its midpoint in the direction in which the line to which it belongs deviates. This effect is much stronger in the illusion of the oblique intercept described by Schilder and Wechsler.[2] In this illusion, if the spacing of the parallel lines is progressively varied and the transverse line is given different slopes with respect to the parallel system of lines (Fig. 2), the deviation of segments, determined on this transverse line by the lines of the field, appears to be an inverse function of the length of the segments and of the angles they form with the lines of the field. The smaller the segments and the angles, the greater the deviations. The same relations seem to hold for the segments determined on the parallels by the cutting lines of the field in Zöllner's and other related illusions.

The deviation of the segments of the parallel lines cannot account for the characteristics observed in *illusions of direction* since, in such illusions, different phe-

* Received for prior publication May 23, 1960.

[1] R. S. Woodworth, *Experimental Psychology,* 1938, 645.

[2] P. Schilder and D. Wechsler, The illusion of the oblique intercept, *J. exp. Psychol.,* 19, 1936, 747-757.

nomena are encountered. In Zöllner's illusion, the two parallel lines appear straight but tilting toward or away from each other; in Wundt's figure, they curve in opposite directions at the sides of the figure.

Another important factor that seems to contribute to these illusions is the *Gestalt principle of direction.* According to this principle the common direction followed by a series of stimulus-items is the basis of grouping and determines the configuration seen. Wertheimer used dots, letters, and small circles to illustrate this princi-

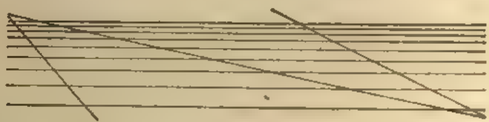


FIG. 2. MODIFIED ILLUSION OF THE OBLIQUE INTERCEPT

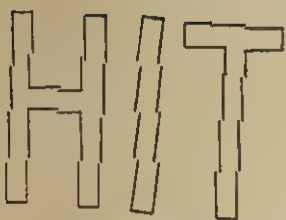


FIG. 3. ILLUSTRATION OF THE GESTALT PRINCIPLE OF DIRECTION EXTENDED TO SEGMENTS OF LINES

The letters look roughly straight but tilting in the direction of the stimulus-items that compose each of them. Here we have the same type of effect as that observed in Zöllner's Illusion

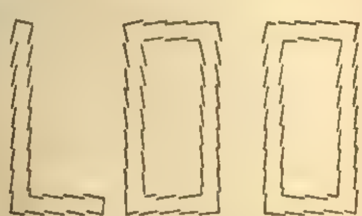


FIG. 4. ANOTHER ILLUSTRATION OF THE GESTALT PRINCIPLE OF DIRECTION EXTENDED TO SEGMENTS OF LINES

Here we have curving effects as in Wundt's and Hering's Figures

ple.[3] Using segments of line as stimulus-items, we can see in Figs. 3 and 4 that the configurations seen are not only determined by the common direction of the segments (their midpoints all falling along straight vertical and horizontal lines) but also by the direction the stimulus-items have relative to one another. In Fig. 3, the direction the stimulus-items have relative to one another is uniform; the letters appear roughly straight but tilting in the direction of the segments that compose each of them. In Fig. 4, this direction is progressively changing from segment to segment; the letters appear to be curving.

In Zöllner's and other related illusions, if the segments determined on the paral-

[3] Max Wertheimer, Untersuchungen zur Lehre von der Gestalt, *Psychol. Forsch.*, 4, 1923, 301-350. (English translation in W. D. Ellis, *A Source Book of Gestalt Psychology*, 1938, 71-88.)

lels by the cutting lines of the field are perceived as deviating, *i.e.* turning around their midpoints as an inverse function of their length and of the acute angles they form with the lines of the field, the parallels present a series of stimulus-items that are organized according to the Gestalt principle of direction extended to segments of lines, as illustrated in Figs. 3 and 4. The principle of pregnance, or good-figure, accounts for the fact that the observer overlooks the irregularities, and also for his tendency to see the characteristics exhibited by these geometrical-optical illusions of direction as fully expressed as the conditions allow.

Unless the segments are perceived as deviating in equal or different amounts from the parallels to which they belong, the principles of direction and pregnance do not, however, operate. Are such deviations actually perceived?

Experiments I and II were conducted to test the following hypotheses. In Zöllner's and other related figures, the deviation of the segments determined on the parallel lines by the transverse lines is: (1) an inverse function of the length of these segments; and (2) an inverse function of the size of the acute angles they form with the intersecting lines of the field.

Experiment I: Length of Segments

Method. The test-objects used were straight lines and sections of Zöllner's illusion (see Fig. 5). Twenty stimulus-objects, 10 test-lines and 10 complex patterns, were used in testing Hypothesis 1.

The different sizes given to the 10 sections of Zöllner's figure in Experiment I were determined by the length of the oblique segment between the two parallels

Fig. 5. One Test-Line and One Section of Zöllner's Figure Used in Experiment I

Both the test-line and the oblique segment in the complex pattern have a slant of 15° with respect to the horizontal

in each pattern. The different lengths adopted were: 1½, 2½, 5, 7½, 10, 16, 22, 28, 34, and 40 mm. With these, the horizontal parallels were given the following dimensions: 5, 8, 10, 15, 21, 28, 37, 51, 57, and 65 mm., respectively, to keep the proportions between the parts of the different test-objects approximately constant. The oblique line in each pattern was given an angle of 15° with respect to the horizontal parallels. Ten straight lines having the slant and length corresponding to the oblique segment in each of the 10 patterns mentioned above were used to determine the point of subjective equality in the absence of distorting factors. All the test-objects were drawn in India ink with a Hunt point #107, on 20 pieces of Bristol board, 8½ in. on each side. Each test-object was so located on the board that the midpoint of the oblique line in each pattern and the corresponding straight lines fall exactly on the point around which the movable part of the instrument described below revolved.

The apparatus (Fig. 6) consisted of the following: (a) a fixed center section, 13 in. in diameter; (b) a frame fastened on it to receive test-objects and secure

them tightly in place; (c) a middle circular movable section, 22 in. outside diameter and 14 in. inside diameter, on which two pointers were secured. The pointers were long enough to reach the scale drawn on the outer edge of the third circular section. Each pointer could be moved forward or backward and locked at a given position. *S* could move the middle section by turning a knob attached to a small wheel touching this movable section. This arrangement allowed him easily to line up the pointers with the test-objects; (d) an outer circular section, 71⅝ in. in diameter with a scale in degrees drawn at its edge; (e) a large screen placed be-

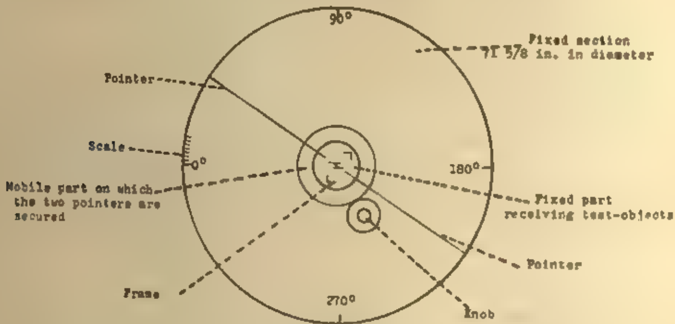


FIG. 6. SCHEMATIC DIAGRAM OF THE APPARATUS

tween *S* and the apparatus to prevent him from reading the alignments made; and (f) a viewing tube, 22 in. from the center of the apparatus. Illumination of the test-objects was provided by a 100-w. bulb placed 16 in. from the test-object and 17 in. below the viewing tube. The general illumination of the testing room was kept constant.

Subjects. Ten men, college students ranging in age from 20 to 23 yrs., served as the *S*s.

Procedure. *S* sat in front of the apparatus and, using his preferred eye, observed through the viewing tube. His task was to line up the pointers with the test-lines and the oblique segments between the parallels in the complex patterns presented to him.

Instructions. At the beginning of the first experimental session, *S* was given the following instructions:

The following test-objects will be presented to you through this viewing tube. At the signal 'Ready' you are to look through the viewing tube and align these two pointers as perfectly as you can with these test-lines. You will do so by turning this knob. Take all the time you need to make your alignments as accurately as you can and let me know when you have finished. Please, do not again look at the test-object after telling me you have finished."

The different objects in question were successively shown to *S* by *E* as required by these instructions. Following the instructions, *S* was immediately given three practice trials with three different test-lines to familiarize him with the testing-situation. Three practice trials were also given *S* the first time he was presented

with the complex patterns. For the last three sessions, shorter instructions were given as follows:

> Now, you will be presented with these test-objects [which were shown *S*]. Your task will be the same as before. Take all the time you need to make your alignments as accurately as you can and let me know when you have finished. Please, do not look again at the test-object after telling me you have finished.

With the presentation of a test-line, the pointers were moved forward or backward by *E* according to the size of the test-object. The standard distances the pointers were moved from both ends of the test-lines for each size (1½, 2½, 5,

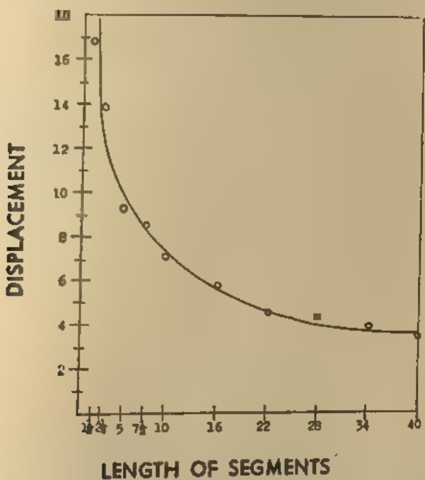


FIG. 7. COURSE OF DISPLACEMENT IN DEGREES AND MINUTES AS A FUNCTION OF LENGTH OF SEGMENTS

Each circle represents an average of 100 determinations, 10 from every *S*.

7½, 10, 16, 22, 28, 34 and 40 mm.) were 5, 6, 8, 10, 14, 16, 18, 23, 25, and 27 mm., respectively, to keep the distance between the pointers and both ends of each test-object proportionally constant. The same standard distances were also used with the 10 complex patterns.

Before every trial, the pointers were turned randomly higher or lower to the actual inclination of the test-objects (15°) and at varying distances from that point to avoid giving *S* positional tendencies, *i.e.* tendencies to turn the knob a certain amount, etc.

To collect the data required for the test of Hypothesis 1, four testing-sessions of 1¾ hr. each were required of every *S*. At each session, two or three series of the 10 test-lines were first randomly presented to *S* (first session, 2 series; second session, 3 series, etc.). After a 5-min. rest-interval, 2 or 3 series of the 10 sections of Zöllner's illusion were randomly shown to him (first session, 3 series; second session, 2 series, etc.). A rest-period of 1 min. 45 sec. was always given *S* between successive trials. The frequency and length of these rest-intervals were to prevent fatigue and cumulative effects of satiation.

During this experiment each test-object was presented 10 times to every *S*; 5 ascending and 5 descending readings were taken with every pattern and every *S*. *E* recorded in degrees and minutes each alignment made by *S*.

Results. The data obtained under the conditions of Experiment I appear in Table I and Fig. 7. The general nature of the relationship between displacement and the length of segments appears to be curvilinear. Displacement is at a maximum (16° 51′) with the smallest segment (1½ mm.) and at a minimum (3° 29′) with the largest pattern (40 mm.). Here the measure of displacement is the deviation of the mean test-settings of the complex patterns from a point of subjective equality given by the mean alignment-settings of the test-lines. The data yield a negative rho of 1.00 between displacement and the length of segments.

TABLE I
MEAN TEST-SETTINGS AND DEVIATIONS
(Every mean is an average of 100 trials)

Independent variable (mm.)	Test-lines	Complex patterns	Mean deviations
1½	16° 5′	32° 56′	16° 51′
2½	14° 56′	28° 51′	13° 55′
5	15° 10′	24° 24′	9° 14′
7½	14° 54′	23° 19′	8° 25′
10	15°	22° 10′	7° 10′
16	15° 4′	20° 47′	5° 43′
22	14° 57′	19° 34′	4° 30′
28	15° 7′	19° 22′	4° 15′
34	14° 44′	18° 28′	3° 44′
40	14° 48′	18° 17′	3° 29′

EXPERIMENT II: SIZE OF ACUTE ANGLES

Procedure. The general experimental procedure and the *S*s were the same as in Experiment I. Twelve test-objects were used: 1 test-line and 11 sections of Zöllner's illusion. The pointers were locked at a distance of 10 mm. from the ends of the test-line and were maintained in the same position throughout the entire experiment.

The test-line as well as each segment between the parallels in the 11 patterns were 5 mm. long, with a slant of 35° with respect to the horizontal. The parallel lines were 13 mm. long in every pattern. These parallel lines were given the following angles with respect to the oblique line from pattern to pattern: 10 ,15, 20, 25, 30, 35, 40, 45, 50, 60, and 70°.

To collect the data required for the test of Hypothesis 2, two testing periods of 2½ hr. each were given every *S*. At the beginning of the two sessions, *S* was first presented with the test-line for 5 trials; then, in random order, with 5 series of the 11 complex patterns. A 5-min. rest-period was given after the completion of the second series. At the beginning of the first testing-period, 3 practice trials were given *S* with the test-line to familiarize him with the new slope (35°) of the test-

objects. Three practice trials were also given with the presentation of the complex patterns. Five ascending and five descending readings were taken with every *S* and every pattern.

Results. The results shown in Table II and Fig. 8 portray the general relationship found between displacement of the segments and the size of the angles they form with the lines passing at their ends. The function appears to be rectilinear. Displacement is at a maximum (13° 14′) with the

TABLE II

MEAN TEST-SETTINGS AND DEVIATIONS

Each Mean is an Average of 100 Trials. Mean test-settings of test-line, 34° 58′.

Independent variable	Complex patterns	Mean deviations
10°	48° 12′	13° 14′
15°	47° 12′	12° 14′
20°	45° 33′	10° 35′
25°	45° 10′	10° 12′
30°	43° 29′	8° 31′
35°	42° 26′	7° 28′
40°	41° 20′	6° 22′
45°	40° 12′	5° 14′
50°	39° 40′	4° 42′
60°	37° 18′	2° 20′
70°	36° 42′	1° 44′

smallest angle (10°) and decreases steadily with the successive increases of angles. The data also yield a negative rho of 1.00 between the dependent and independent variables.

DISCUSSION

The results of both experiments support the respective hypotheses under test. In Zöllner's figure and other related illusions of direction, each segment of the parallel lines was found to be deviating as an inverse function of its length and of the acute angles it forms with the cutting lines of the field. The data yield a negative rho of 1.00 between displacement and the length of the segments, apparently curvilinear, and a negative rho of 1.00 between displacement and the size of the angles given to the cutting lines, apparently rectilinear.

The results of both experiments also seem to validate, within the limits of the research design, the Gestalt principle of direction evoked at the beginning to account for the direction and form the parallel lines take in these illusions of direction. According to the Gestalt principle of direction extended to segments of line, the direction and form the parallel lines should take in these illusions are not only a function of the common direction of their segments (all falling along perfectly straight vertical or

horizontal lines), but also a function of the direction the segments have relative to one another. In Zöllner's figure, the segments of the parallel lines have a common direction; all their midpoints fall along straight vertical lines; and the direction the segments have relative to one another is constant since they all have the same length and form equal angles with the transverse lines of the field. In this figure, two straight lines are seen tilting slightly toward each other according to the direction the segments have in each parallel. Thus, the principle of direction applies here. In

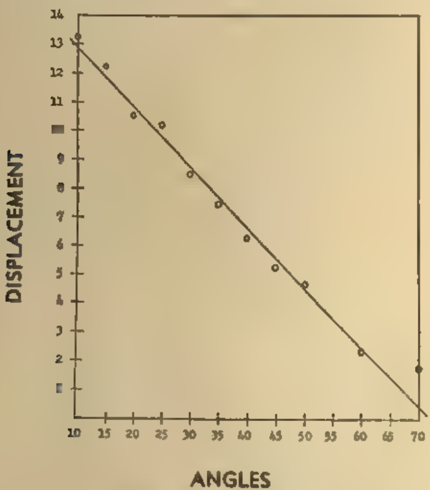


FIG. 8. COURSE OF DISPLACEMENT IN DEGREES AND MINUTES AS A FUNCTION OF ANGLES

Each circle represents an average of 100 alignments, 10 from every *S*.

Wundt's figure, the segments of each parallel line have also a common direction; all their midpoints lie along perfectly straight horizontal lines. The direction the segments have relative to one another from the middle of the figure to the sides changes, however, progressively from segment to segment as a result of their slightly varying lengths and of the progressively smaller and smaller angles they form with the transverse lines of the field. The two parallel lines are seen as curving away from each other at both ends according to the direction the segments have in each parallel. The principle of direction applies here as it applies to Hering's figure and to some other related figures.

Other evidence of the relation of geometrical factors to the principle of direction is the fact that the curving effects observed in Wundt's and

Hering's figures can be produced with Zöllner's figure alone by progressively increasing the length of the segments determined on the parallel lines by the cutting lines of the field, or by giving the transverse lines of the field progressively smaller or larger angles with respect to the parallels. A combination of these two variables in the same figure gives an expected stronger curving effect.

The principle of pregnance or good-figure, in the optical illusions of direction, would account for the fact that we overlook the irregularities created by the deviations of the segments of the parallels and see these lines as better lines than they actually appear.

Summary

Although displacement of whole parallels in illusions of direction depends upon the Gestalt principles of direction and pregnance, it is questionable whether the Gestalt principle of direction would operate unless deviations of the segments determined on these parallels by the transverse lines of the field are reacted to by *S*.

It was hypothesized that in Zöllner's illusions and other related figures, the deviation of the segments determined on the parallel lines by the transverse lines is (1) an inverse function of the length of these segments, and (2) an inverse function of the size of the acute angles they form with the intersecting lines. The results of the experiment supported these hypotheses.

CONFIDENCE IN THE RECOGNITION AND REPRODUCTION OF WORDS DIFFICULT TO SPELL

By PAULINE AUSTIN ADAMS, Menlo Park, California and
JOE K. ADAMS, Palo Alto Medical Research Foundation

It has long been known that, under a wide range of conditions, recognition yields higher performance-scores than reproduction. Apparently, however, the question of the *Ss'* confidence in the accuracy of their performances under these two conditions has not been explored. The present study attempts to answer a number of questions relating to accuracy and confidence in the recognition and reproduction of a particular type of material; namely, words that are difficult to spell.

Granting the generalization that recognition yields higher performance-scores than reproduction, our inquiry will be directed to a comparison of the *appropriateness* of the *Ss'* confidence in the two situations. Is confidence in recognized items more or less realistic than confidence in reproduced items? Do *Ss* have more appropriate confidence when they are reproducing items than when they are recognizing them? A comparison of *S*'s confidence in two different situations can be made only if the scale of confidence he uses has the same objective meaning in the two cases. In the present experiment a scale of confidence defined in terms of probabilities or expected percentages was used.[1] This scale enables us to compare the appropriateness of *S*'s confidence in recognition and reproduciton and to answer the question whether reproduction, with a certain degree of confidence or recognition, with that same degree, is more likely to be correct.

In the present experiment, in addition to indicating their confidence in the accuracy of their performances, the *Ss* rated their familiarity with the meanings of the words. This latter procedure permits us to examine the accuracy of performance and subjective confidence as a function of word familiarity.

METHOD

One hundred stimulus-words having at least three plausible incorrect spellings, *i.e.* spellings frequently used for the same sounds (fiberus, fiberous, fibreous, fibrous) were selected from the Thorndike-Lorge word list.[2] The words were arranged alpha-

* Received for publication July 28, 1959.

[1] J. K. Adams, A confidence-scale defined in terms of expected percentages, this JOURNAL, 70, 1957, 432-436.

[2] E. L. Thorndike and Irving Lorge, *The Teacher's Word-Book of 30,000 Words*, 1944.

betically and assigned to two lists in alternation. The words within each of these lists were rearranged in random order into 5 blocks of 10 words each. Two additional orders for each of the two lists were constructed by rearranging the words in the original blocks. In the procedure used for Recognition, the correct spelling and the three alternative versions were printed in random order on *S*'s record-sheets and he was asked to indicate his confidence after circling the version which he believed to be correct. In the procedure used for Reproduction, *S* indicated his confidence after writing the word. These judgments, in both procedures, were followed by *S*'s report of his familiarity with the stimulus-word. For this purpose a five-point scale was used: very familiar, familiar, slightly familiar, unfamiliar, and unknown. This scale, to aid *S*, was printed on record and blank sheets supplied him. In both procedures, *E* pronounced the words at the rate of one every 10 sec. The confidence-scale which *S* used to indicate his estimate of the probability that he had circled the correct version or had spelled the word correctly, was defined as follows. Under the Recognitive procedure, *S* used a number between 25, which would indicate that he felt that his choice was a pure guess, and 100, which would indicate that he was certain that he had circled the correctly spelled version. Under the Reproductive procedure the confidence-scale ran from 0, which would indicate that *S* believed that there was no chance at all that he had spelled the word correctly, to 100, which would indicate complete certainty in the accuracy of the reproduction. In the Recognitive procedure where there were four alternative versions the point of pure guess was at 25 since by chance alone *S* could be correct in 25% of his judgments. If this scale were used appropriately, then all of the words opposite which *S* wrote 100 should be correct; approximately 80% of the words opposite which *S* wrote 80 should be correct; approximately 50% of the words opposite which 50 was written should be correct, and so forth. He was told that this was the criterion for correct usage of the scale, and also that he was to use only multiples of 5 in indicating his confidence. The various points on the familiarity-scale were defined in terms of *S*'s knowledge of the meaning of the words: 'very familiar' meant that *S* was certain of the meaning of the word; 'familiar,' that he was quite sure of the meaning of the word; 'slightly familiar' that he thought that he knew the meaning of the word but was not sure; 'unfamiliar,' that he did not think that he knew the meaning of the word but had heard the word before; and 'unknown,' that he was certain that he did not know the meaning of the word or did not think that he had heard the word before.

There were two experimental groups, a Recognitive group and a Reproductive group. The test of reproduction, however was followed by one of recognition. There were, therefore, three experimental conditions: Recognition, Reproduction, and Reproduction-Recognition, *i.e.* recognition preceded by reproduction.

Subjects. One hundred undergraduate students served as *S*s. The experiment was conducted with groups ranging in size from 2 to 7 *S*s. There were 50 *S*s in each experimental situation: 25 receiving each of the two word-lists. The *S*s under the Reproductive procedure also provided the recognitive data for the Reproductive-Recognitive procedure.

Scoring. In addition to percentage correct, a mean confidence-score was obtained for each *S*. As the measure of appropriateness-of-confidence a discrepancy-score was computed. As explained above, appropriate usage of the confidence-scale would require a close correspondence between a given point on the confidence-scale and

the percentage of judgments given with that confidence which were correct, taking into account, of course, variability expected from sampling. Correct usage of the confidence-scale would also result in a close correspondence between the mean confidence for all judgments and the percentage of correct judgments. The discrepancy between these two measures (mean confidence and percentage correct) represents the appropriateness of the *S*s' confidence. Perfect appropriateness would yield scores of zero discrepancy; over-confidence, positive scores (mean confidence exceeding percentage correct); and under-confidence, negative scores (percentage correct being greater than mean confidence). The discrepancy-score for each *S* was calculated by subtracting percentage correct from mean confidence, thus the score for any given *S* might be positive or negative in value.

Results

In representing the accuracy for each *group* of *S*s two discrepancy-scores were obtained: an absolute discrepancy-score which was based on the sum of the absolute values of *S*'s scores; and an algebraic discrepancy-score, which used the algebraic sign of *S*'s scores. The tendency for the group as a whole to be over- or under-confident is indicated by the extent to which these two scores are similar in magnitude as well as by the size of the algebraic discrepancy-score. Table I summarizes these measures for each of the experimental conditions, and the *t*-values resulting from the comparisons of the various conditions on the confidence- and accuracy-measures.

Although the results of the two Recognitive procedures are not significantly different from each other in any of these comparisons, the Reproductive condition differs signficantly ($p < 0.01$) from both of these conditions in all comparisons. Under the Reproductive procedure, the *S*s have lower confidence than in the Recognitive. In the Reproductive as was to be expected, fewer of the *S*s' responses were correct.

The answer to the main problem toward which this study was directed is found in the comparisons of the discrepancy-scores. Appropriateness-of-confidence is greater for Recognition than Reproduction. Under the conditions of our experiment, the *S*s were over-confident in both Recognition and Reproduction, but they were over-confident to a greater degree when reproducing items than when recognizing them. This result is revealed in the significantly higher discrepancy-scores for Reproduction than for Recognition. Confidence was more appropriate for recognized items at all points on the confidence-scale. This result is shown in Fig. 1. There is a difference of about 10 points (in percentage correct) all along the scale. An item recognized with a given degree of confidence had 10% greater probability of being right than an item reproduced with that same confidence.

The direction of this result cannot be due to the different scales for the two procedures (the scale for Recognition running from 25, and the scale for Reproduc-

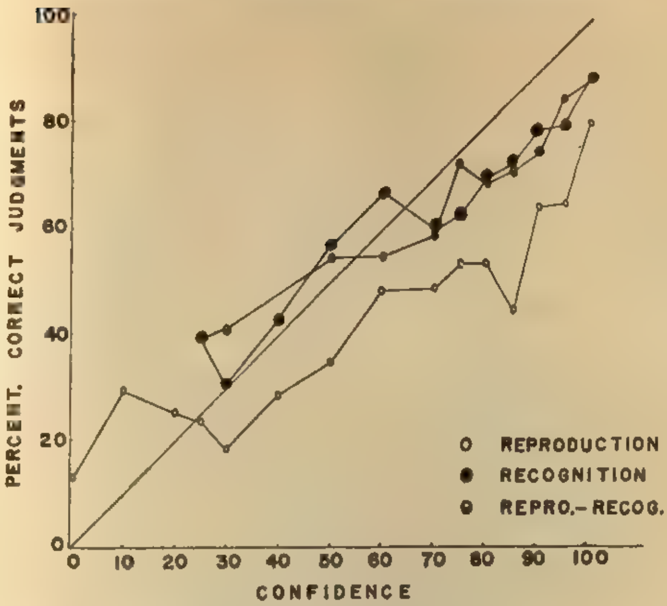


FIG. 1. CONFIDENCE-FUNCTIONS FOR REPRODUCTION, RECOGNITION, AND REPRODUCTION-RECOGNITION

TABLE I

MEANS FOR THE CONFIDENCE- AND ACCURACY-MEASURES AND SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE MEANS

Procedure	Means for the confidence and accuracy scores: Mean confidence	% Correct	Discrepancy: absolute	Discrepancy: algebraic
Reproductive	72.38	57.28	16.60	15.10
Recognitive	81.18	75.20	9.03	5.98
Reprod.-Recog.	84.11	76.88	9.57	7.23

Differences in means with t-values for each difference

Scores	Reprod. minus Recog.		Reprod. minus Reprod.-Recog.		Recog. minus Reprod.-Recog.	
	diff.	t	diff.	t†	diff.	t
Mean confid.	− 8.80	3.65*	−11.73	7.72*	−2.93	1.51
% correct	−17.92	5.64*	−19.60	11.81*	−1.68	0.66
Discrepancy:						
absolute	7.57	4.05*	7.03	5.25*	−0.54	0.35
algebraic	9.12	3.93*	7.87	4.95*	−1.25	0.61

* Significant at the 1% level.

† t-values are based on paired scores as the same *S*s are involved.

tion from 0) as the confidences for Recognition would have been the same or lower if the 0–100 scale had been used, whereas the confidences for Reproduction would have been the same or higher if the 25–100 scale had been used for that group.

The graph in terms of under- and over-confidence at the various points on the confidence-scale can be described as follows. The Reproductive group tended to be under-confident by about 10 points at confidences 0–20, to be realistic at 25, to be over-confident by about 10 points from 30–60, and to be over-confident by about 20 points from 65–100. The Recognitive group was under-confident by about 15 points at confidence 25, realistic from 30–60, and over-confident by about 10 points from 70–100.

The use of a confidence-scale defined in terms of probabilities permits us to compare *S*s' appropriateness-of-confidence independently of their percentage of correct judgments. The above finding of a higher discrepancy-score (in this case, greater over-confidence) suggests that in the Reproductive condition although *S*s are less confident, their confidence was nevertheless relatively higher than their percentage correct. It might be argued, however, that if the *S*s are in general over-confident, in a situation (recognition) where their percentage correct tends to be greater, discrepancy-scores will be lower as a simple mathmatical result of this type of ceiling-effect. To ascertain whether the finding of a higher discrepancy-score in the Reproductive condition could be attributed to the smaller percentage of correct judgments in this procedure, an analysis was performed on the results of pairs of *S*s matched on the basis of percentage of correct responses. The *S*s in the Reproductive and the Recognitive procedures were arranged in rank order according to their percentage of correct judgments. They were then so matched that the mean percentage correct was the same for the two conditions. In 11 of the resulting 20 pairs, the paired *S*s had identical numbers of correct responses, while there was a difference of one item for 6 pairs, and 2 items for 3 pairs. A t-test was then carried out on the differences between the algebraic discrepancy-scores for the pairs of *S*s. The t-value of 2.68 ($p < 0.02$) obtained, suggests that the finding of a significantly higher discrepancy-score in Reproduction should not be attributed to the smaller number of correct judgments in this condition.

There are at least two factors accounting for a higher performance-score in recognition than in reproduction. One is that in recognition *S* is provided with cues—part of the process of memory is accomplished for him. The second factor is simply that in recognition sheer guessing results in an appreciable percentage of correct answers. There may be additional factors of a less obvious nature, *e.g.* motivational or emotional processes. In any

case, it seems reasonable that whatever factors account for the higher performance-scores also account for the increase in confidence—or at least that there is considerable overlap between the two sets of factors. Under our conditions, confidence did not increase as much as performance; thus discrepancy declined. This result raises the more general question: granted that an increase in information (or cues) increases both performance and confidence, is there in general a decrease in the discrepancy between them; in other words, do they increase differentially so as to narrow the gap? We doubt very much that such a generalization is valid. It seems more plausible that information can be added in such a way as to increase performance faster than confidence or contrariwise.

Correlations between accuracy and confidence. There are three questions that may be asked from our data concerning the relationships between levels of confidence, percentage correct, and appropriateness-of-confidence. (1) Do *S*s with relatively high confidence also tend to be correct on a larger percentage of their judgments; (2) is there a relationship between confidence and appropriateness-of-confidence; and (3) do *S*s with a larger number of correct responses also tend to be more appropriate in their confidence? These questions can be answered by an examination of the correlation-ratios among the three measures: mean confidence for all items; percentage correct; and absolute discrepancy-score. Scatter-plots indicated that the use of r would be justified. For each of these comparisons a correlation coefficient of 0.37 can be considered to be significant at the 5% level. Table II summarizes these correlations.

For all three of the experimental conditions, a higher mean confidence is correlated with a higher percentage of correct responses. It is important to note that, although confidence measures and percentage correct are highly interrelated,[8] only the percentage correct correlates with discrepancy; the correlation between mean confidence and discrepancy is negligible. The correlation between percentage correct and discrepancy is negative indicating that *S*s with more accurate performances are also more appropriate in their confidence-judgments.

Individual consistency in the use of the confidence-scale. When we examine the results for the same *S*s in the Reproductive and Reproductive-Recognitive procedures we find significant individual consistency in the use

[8] These correlations are higher than those usually found between confidence and performance (cf. D. M. Johnson, *The Psychology of Thought and Judgment*, 1955). A smaller than usual variability in meaning (of the confidence-scale) from *S* to *S* could account for the higher correlations.

of the confidence-scale. All the confidence- and accuracy-scores are correlated to a significant degree. This result is summarized in Table III.

Distribution of confidence-judgments. Distributions of confidence-judgments over the range 25–100 for the three procedures were remarkably similar, the only exception being the relatively greater use of 25 in the Recognitive procedure. Similarly, we find a peak in the distribution curve at 0 for the Reproductive procedure. For all conditions about 40% of the *S*s' judgments were given with 100% confidence. Peaks in the distribution-

TABLE II
CORRELATIONS BETWEEN THE CONFIDENCE AND ACCURACY MEASURES

	Reproduction	Recognition	Reprod.-Recog.
Mean confidence vs. % correct	0.88	0.65	0.60
Mean confidence vs. absolute discrepancy	0.18	0.11	0.11
% correct vs. absolute discrepancy	−0.32	−0.53	−0.57

TABLE III
CORRELATIONS BETWEEN SCORES IN THE REPRODUCTION AND REPRODUCTION-RECOGNITION PROCEDURES

Mean confidence	% correct	Absolute discrepancy
0.52	0.77	0.49

curves occur at confidences of 50%, 80%, and 90%. Approximately 8% of the *S*s' judgments were made with these confidences. In all other cases, except for 75% confidence the confidence-points between multiples of 10 are used less frequently than the multiples of 10.

Relations between confidence, accuracy, and familiarity. Examples of words which were rated as very familiar, and least familiar are given in Table IV. This table also presents examples of words which were correctly recognized and reproduced by a large majority of the *S*s and also by very few of them. Most of the words, as the *S*s indicated, were quite familiar. Distributions for the three procedures were very similar in the percentage of judgments given in each of the five familiarity-classes. From the most to the least familiar these percentages (averages for the three procedures) were: 67.29, 18.77, 6.88, 3.51, and 3.55. The words selected for this study represented such a narrow range of the Thorndike-Lorge word-frequency counts (range on the L count 0–531, median 11) that it would not be reasonable to try to relate word-frequency to our confidence- and accuracy-measures. Also, there does not appear to be a very close rela-

tionship between word-frequency and the individual *S*'s ratings of the familiarity of the words. Our concern will, therefore, be with the relations between the *Ss*' ratings of familiarity, the percentage of their judgments that are correct for each word-familiarity classification and the mean confidence, and appropriateness-of-confidence associated with the words in each familiarity-class. For each *S* the mean confidence for all responses as well as the percentage of correct responses for words given each familiarity rating were determined. In addition, absolute discrepancy-scores were

TABLE IV

EXAMPLES OF FAMILIAR AND UNFAMILIAR, AND EASY AND DIFFICULT STIMULUS-WORDS

(A) Examples of most familiar and least familiar words. Entries indicate the number of *Ss* using each familiarity category. (a) Reproduction, (b) Recognition, (c) Reproduction-Recognition.

Words	Familiarity rating 1			2			3			4			5		
	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
Familiar															
commission	21	24	18	4	1	7									
procedure	25	24	24	0	1	1									
Unfamiliar															
fissure	8	1	9	1	11	4	2	1	4	4	6	2	10	1	6
appurtenance	7	1	6	1	4	2	9	8	5	4	9	7	4	3	5

(B) Examples of easy and difficult words. Entries indicate the number of *Ss* correctly reproducing or recognizing each word

Words	Reproduction	Recognition	Reprod.-Recog.
Easy			
masquerade	20	25	25
racketeer	21	25	24
Difficult			
balustrade	2	8	6
paraphernalia	3	1	6

obtained for each familiarity-class from the group-results. For all three procedures the mean confidence decreases approximately linearly with familiarity when we plot the familiarity-points at equal intervals. The curves for the three conditions are approximately parallel. The mean confidence for words classified as 'very familiar' is about 85; for the other familiarity-classifications, the mean confidences are approximately 75, 62, 48, and 30. The percentage of correct responses in each familiarity-class does not, however, demonstrate such a precise relationship; in addition there are differences between the three experimental procedures in this relationship. In the Reproductive procedure there are no significant differences in percentage

correct between words in the two most familiar word-classes, nor is the difference in percentage correct significant for the words in the two least familiar-word classes. The differences in the middle of the familiarity-scale are significant (between the second and third points on the scale, $t = 2.36$, $p < 0.05$; and the third and fourth, $t = 3.40$, $p < 0.01$). In the Recognitive procedure the relation between percentage correct and familiarity is linear though the only significant difference in percentage correct obtained between consecutive classes was between the two most familiar word classes ($t = 2.53$, $p < 0.05$). The relation in the Reproductive-Recognitive procedure is irregular, suggesting that the procedure of first reproducing the words minimizes the relation between apparent familiarity and accuracy. There is no statistically significant relationship between familiarity and appropriateness-of-confidence. There is, however an indication of an increase in discrepancy with increasing familiarity; in other words, with increasing confidence increased faster than percentage correct.

Conclusions

(1) Under the conditions of our experiment, confidence in recognized items was more appropriate than confidence in reproduced items. This was true at all points on the confidence-scale. Both conditions yielded overconfidence at the upper end of the confidence-scale.

(2) Recognition following reproduction yielded approximately the same percentage correct, mean confidence, and appropriateness-of-confidence as the initial Recognitive procedure.

(3) Familiarity as rated by the *Ss* was found to be highly related to confidence-judgments and to percentage correct. Overconfidence was found to increase with increasing familiarity. This increase, however, was not statistically significant.

(4) Distributions of confidence-judgments for the three procedures were remarkably similar.

(5) Correlations of from 0.60 to 0.88 were found between mean confidence and percentage correct.

AN EXPERIMENTAL ANALYSIS OF SET: VARIABLES INFLUENCING THE IDENTIFICATION OF AMBIGUOUS, VISUAL STIMULUS-OBJECTS

By EUGENE R. LONG, University of North Carolina, L. STARLING REID, and RICHARD H. HENNEMAN, University of Virginia

Gibson has pointed out that the concept of set in its relation to perception has been widely investigated.[1] For the most part the literature agrees that accuracy in describing or identifying impoverished stimuli can be altered by the establishment of particular sets. Opinions have long differed, however, as to the way in which these changes are effected. For example, the results of Külpe and Bryan; Yokoyama; Chapman; Postman and Bruner; Neisser; Krulee, Podell, and Ronco; Green and Anderson; Ross, Yarczower, and Williams; and Hoisington and Spencer support an hypothesis of perceptual selectivity or sensitization.[2] This hypothesis is based on the finding that *Ss* usually best identified or described those characteristics of the stimulus-field about which they had received the most specific instructions, and thus for which they were best set prior to the presentation of the stimulus. The data of others, *e.g.* Lawrence and Coles and Lawrence and LaBerge, do not support this hypothesis.[3] Instead they strongly suggest that set influences retention and response rather than the reception of a stimulus.

* Received for publication June 4, 1959. This study is based upon a modification of a Technical Report written under Contract W 33 (038)-ac-21269 between the University of Virginia and the USAF, Psychological Branch, Aero-Medical Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio.

[1] J. J. Gibson, A critical review of the concept of set in contemporary experimental psychology, *Psychol. Rev.*, 38, 1941, 781-817.

[2] Oswald Külpe, Versuche über Abstraktion, *Ber. I. Kongr. exp. Psychol.*, 1904, 56-68; E. G. Boring, Attribute and sensation, this JOURNAL, 35, 1924, 301-304; D. W. Chapman, Relative effects of determinate and indeterminate Aufgaben, this JOURNAL, 44, 1932, 163-174; L. J. Postman and J. S. Bruner, Multiplicity of set as a determiner of behavior, *J. exp. Psychol.*, 39, 1949, 369-377; Ulric Neisser, An experimental distinction between perceptual process and verbal response, *ibid.*, 47, 1954, 399-402; G. K. Krulee, J. E. Podell, and P. G. Ronco, Effect of number of alternatives and set on the visual discrimination of numerals, *ibid.*, 48, 1954, 75-80; B. F. Green and L. K. Anderson, Color coding in a visual search task, *ibid.*, 51, 1956, 19-24; Sherman Ross, Matthew Yarczower, and G. M. Williams, Recognitive thresholds for words as a function of set and similarity, this JOURNAL, 69, 1956, 82-86; L. B. Hoisington and Carol Spencer, Specific set and the perception of subliminal material, this JOURNAL, 71, 1958, 263-269.

[3] D. H. Lawrence and G. R. Coles, Accuracy of recognition with alternatives before and after the stimulus, *J. exp. Psychol.*, 47, 1954, 208-214; D. H. Lawrence and D. L. LaBerge, Relationship between recognition accuracy and order of reporting stimulus dimensions, *ibid.*, 51, 1956, 12-18.

In an earlier article the present authors suggested a model which would account for the apparently contradictory results.[4] Similar models have also been proposed by Garner, Hake, and Eriksen and by Kanfer.[5] Specifically, we postulated that complex perceptual behavior is composed of a series or chain of responses that fall into three classes, entailing (1) stimulus-discrimination, (2) identification or recognition, and (3) instrumental responding. Further, the stimuli that have been employed in most perceptual situations are ambiguous, *i.e.* are capable of eliciting more than one response in a particular class of responses. The ambiguity results from impoverishment, distortion, or the simultaneous presence of other stimuli. Set operates as a response-'limiter.' It increases perceptual accuracy by increasing the probabilities of certain responses and decreasing the probabilities of others in the various response-classes.

This model was designed to account for previous experimental results. In addition it indicates some questions which must yet be answered.

(1) Can set vary in degree, or is it all-or-none in nature?

(2) Are various methods of limiting responses equally effective in increasing perceptual accuracy?

(3) Does type of degradation, *e.g.* omission of part of the stimulus as compared to the addition of other stimuli, influence the effectiveness of set?

(4) Does level of degradation, and thus level of ambiguity, influence the effectiveness of set?

(5) Does set operate in the same manner for all classes of stimuli?

(6) Under what experimental conditions might set operate: (a) to make residual elements of degraded stimuli more discriminable; (b) to make the residual elements more interpretable by supplying or eliminating responses; and (c) to increase correct responding simply by reducing the number of alternative responses.

Question 1 might be answered by varying the number of aternative responses, the population from which they are chosen, and the temporal position of their presentation. Different methods of limiting responses (Question 2) are varying number of alternative responses or presenting more attributes characteristic of the undistorted stimulus. Question 3 might employ omission of part of the stimulus, addition of extraneous stimulus-

[4] E. R. Long, R. H. Henneman, and L. S. Reid, Theoretical considerations and exploratory investigation of "set" as response restriction, *USAF, WADC tech. Rep.*, 53-311, 1953, 1-24.

[5] W. R. Garner, H. W. Hake, and C. W. Eriksen, Operationism and the concept of perception, *Psychol. Rev.*, 63, 1956, 149-159; F. H. Kanfer, Perception: Identification and instrumental activity, *ibid.*, 63, 1956, 317-329.

elements, and the destruction of contours by misfocusing. Question 4 can be investigated by varying level of degradation simultaneously with number of alternative responses. Question 5 might vary the classes of stimuli, *e.g.* words, letters, and pictures, or the sense modality. Question 6 can be answered by comparing the effects of sets given both before and after the stimulus with a set given only after.

I. Variables Influencing the Identification of Ambiguous Visual Stimulus-Objects

The experiments reported in Section I deal primarily with Questions 1, 4, and 6. The remaining questions will be considered later.

Experiment I

The purpose of the first experiment was to determine: (1) whether set can vary in degree; (2) if its effects are related to the level of ambiguity; and (3) under what conditions it augments discrimination, or increases interpretability, or increases correct responses simply by reducing the number of alternatives. The general plan was to present for identification degraded letters of varying degrees of ambiguity. Set (a restriction of response) was manipulated by presenting different numbers of alternatives and different temporal arrangements.

Method. The stimuli were patterns projected on a screen from 35-mm. positive transparencies of degraded single-letter patterns. These had been previously employed in a series of legibility studies.[8] They were prepared by photographing patterns of letters composed of black and white elements. The contours of the letters were degraded by omitting elements of the letters, by adding new elements, or by both omitting some elements and adding others. The projected letters subtended visual angles approximately 6° in height and 4° in width, with the width of strokes being approximately 1°.

The stimulus-letters were scaled according to the frequency of their correct identification in the earlier legibility studies. The assumption was that because of the degree of degradation some patterns were ambiguous and that the most ambiguous patterns were the ones that had been least frequently recognized in the earlier legibility studies. The least ambiguous letters used in this experiment (those letters labeled Lever 1) had been identified approximately 51% of the time. The other two levels of ambiguity and the frequencies of their previous identification were: Level II, 18% and Level III, 4%. Although only 11 letters (A,B,E,K,Q,P,R,

---

[8] For a detailed description of the make-up of these letters and the methods of producing contour-degradation, see Long and Reid, Factors determining the legibility of letters and words printed in 'dot' patterns with pure black and white when the patterns are degraded in varying amounts, *USAF, WADC, tech. Rep.*, 59-22, 1952, 1-23.

S,W,Y,Z) were used, there were three replications of each letter, each employing a different pattern with the result that each level of ambiguity was represented by a total of 33 different slides.

Set was manipulated by allowing *S*s to view small sheets of paper on which were printed lists of 11, 8, 6, or 4 letters—one of which corresponded to the stimulus on that particular trial. This procedure was based on the assumption that as the number of alternative letters was decreased, responses would be more restricted. The distractors (incorrect alternatives) were letters never used as stimuli. Fifteen letters, (C,D,F,G,H,I,J,L,M,N,Q,T,U,V, and Y) were so used. Distractors for a particular stimulus were chosen at random. All letters, however, were used an equal number of times for each particular number of alternatives. Half of the *S*s viewed the sheets of alternatives both before and after presentation of each stimulus and half saw their sheets only after. The sheets for each stimulus at a particular level of ambiguity were combined into booklets, their order being appropriate to the order of the stimuli and the particular temporal arrangement for giving the set.

Restriction of response (number of alternatives) and the temporal arrangements for presenting alternatives were combined as variables in one experimental design. During the first 6 sec. of a trial, *S*s were allowed to look at their sheets of paper. Those *S*s, who were to be set only after the stimulus, were presented a blank sheet of paper. A slide was then projected for 4 sec. after which the *S*s viewed a second sheet for 12 sec. *S*'s task during this period was to indicate on his second sheet of alternatives the letter that had been projected. He then turned over the marked sheet, looked at the next sheet, and then at the next slide. This procedure was repeated 33 times, once for each slide, thus requiring a total time of 12.1 min.

The four degrees of restriction of response, two temporal arrangements for presenting alternatives, and three levels of ambiguity yielded 24 different combinations of experimental conditions. Three *S*s were randomly assigned to each of these combinations, making a total of 72 *S*s. All were undergraduate college men.

Results. The total number of correct identifications made under each of the 24 combinations of experimental conditions (three *S*s $\times$ 33 stimuli per combination) are presented in Table I. An analysis of variance performed on these data indicated that the effects ascribable to two of the three variables were highly significant. Increases in level of ambiguity produced decreases in identification ($F = 64.0$, $p < 0.01$); increases in restriction of response, however, produced increases ($F = 108.5$, $p < 0.01$). Further analysis indicated that this increase was linear in nature. This latter finding is in accord with the data of Frick, Hyman, Archer, and Newbrough and indicates that set is not all-or-none in nature but may vary in degree.[7]

[7] F. C. Frick, Some perceptual problems from the point of view of information theory, in *Current Trends in Information Theory,* 1953; Ray Hyman, Stimulus information as a determinant of reaction time, *J. exp. Psychol.,* 45, 1953, 188-196; E. J. Archer, Identification of visual patterns as a function of information load, *ibid.,* 48, 1954, 313-317; J. R. Newbrough, Interaction between total stimulus information and specific stimulus information in visual recognition, *ibid.,* 55, 1958, 297-301.

Re-analysis of the data after they had been corrected for increases in chance expectancy, using the formula $S = R - [W/(n-1)]$, resulted in some reduction in the advantage gained by reducing the number of alternatives. Identification was still significantly improved by increased restriction of responses, however, and all of the other results were again obtained. It may be concluded, therefore, that set did not increase accuracy of identification by merely reducing the number of alternatives and thus increasing the probability of each.

The average interaction between level of ambiguity and number of

TABLE I

TOTAL NUMBER OF CORRECT IDENTIFICATIONS IN EXPERIMENT I

(Each entry represents the total for three *S*s on 33 stimuli.)

Number of alternatives	Level of ambiguity: I, before and after	I, after	II, before and after	II, after	III, before and after	III, after	Mean
11	64	67	44	44	23	26	44.7
8	67	70	41	51	30	23	47.0
6	71	68	62	59	40	24	54.0
4	80	81	54	45	48	55	60.5
Mean	70.5	71.5	50.3	49.8	35.3	32.0	

alternatives did not attain the 5% level of significance. For the time being the effect of set must, therefore, be assumed to be independent of level of ambiguity.

Finally, viewing alternatives both before and after the presentation of each stimulus resulted in no more identifications than did seeing them only after. These results support the hypothesis that set does not augment the *discriminability* of stimulus-elements.[8] Rather they support the idea that set rendered the residual elements in the distorted stimuli more interpretable by enhancing one of the alternative responses or by eliminating responses capable of being elicited by the same residual elements. This hypothesis resembles the hypotheses suggested previously by Bricker and Chapanis and by Krulee to account for identification under somewhat different experimental conditions.[9]

[8] Lawrence and Coles, *op. cit.*, 1954, 213; Lawrence and LaBerge, *op. cit.*, 1956, 17.

[9] P. D. Bricker and Alphonse Chapanis, Do incorrectly perceived tachistoscopic stimuli convey some information, *Psychol. Rev.*, 60, 1953, 181-188; G. K. Krulee, Some informational aspects of form discrimination, *J. exp. Psychol.*, 55, 1958, 143-149.

Experiment II

In Experiment I letters that were never employed as stimuli were the only ones used as incorrect alternatives. It seems reasonable to assume that set might well be a function of the population from which the alternatives were chosen, as well as of the number. A comparison of the scores of identification obtained with two groups of alternatives drawn from different populations would thus supply a partial answer to Question 2.

The finding that the presentation of alternatives both *before* and *after* each stimulus resulted in no more identifications than the presentation of them only *after* also needs clarification. For example, the number of alternatives was never less than four, and the large numbers of alternatives may have prevented the augmentation of the discrimination of important stimulus elements. If smaller numbers had been employed, the alternatives presented before the stimulus might have resulted in more identifications.

To answer these questions, a second experiment was conducted in which the number of alternatives, the temporal arrangement for presenting the alternatives, and the population of alternatives were manipulated.

Method. Twenty-six stimuli were presented for identification. These stimuli were identical in form and size to those employed in Experiment I. Thirteen different letters (A,B,C,D,E,G,H,I,M,O,P,Q,Z) were used, the two replications of each letter involving the use of a different slide-pattern. One level of ambiguity was used, that being the 5% level.

Set was manipulated by allowing *Ss* to view groups of 8, 4, or 2 alternatives, either both before and after presentation of each stimulus or only after. The groups of alternatives differed in respect to the populations from which they were selected. Half of the *Ss* viewed restricted groups of alternatives, *i.e.* groups in which incorrect alternatives were drawn from a population composed of the 15 letters which were never used as stimuli. The other half viewed unrestricted groups. These were composed of incorrect alternatives drawn from a population of all 26 letters and thus included as incorrect responses letters sometimes used as stimuli. Other procedural characteristics were identical with those of Experiment I.

Six *Ss* were randomly assigned to each of the 12 different combinations of experimental conditions, making a total of 72 *Ss*.

Results. The total number of identifications and the conditions under which they were obtained are presented in Table II. Each entry represents the totals of six *Ss*. An analysis of variance performed on the data indicated that significant effects were produced by all three variables.

Set resulting from the use of restricted groups of alternatives produced more identifications than did that based on the use of unrestricted groups ($F = 17.5$, $p < 0.01$). These results are in accord with those of Miller, Heise, and Lichten, although it should be noted that in the present ex-

periment it was the population of distractors rather than the population of stimuli which was increased.[10] In this regard the present data are more closely akin to those of Krulee, Podell, and Ronco.[11] Here again, however, differences should be noted. The *Ss* in the present experiment were never informed of the different populations from which the incorrect alternatives were selected nor were they told that only 13 of the 26 letters were used as stimuli. Moreover, the *Ss'* post-experimental reports indicated that they were completely unaware of these experimental conditions. The present results, therefore, differ at least in part from those of Krulee, Podell, and Ronco, who found no significant augmentation of identification until their

TABLE II

TOTAL NUMBER OF CORRECT IDENTIFICATIONS IN EXPERIMENT II

(Each entry represents the total number of identifications made by six *S*s on 26 stimuli.)

Number of alternatives	Temporal arrangement: Before and after		After		Mean
	unrestricted	restricted	unrestricted	restricted	
8	30	48	31	40	37.3
4	54	78	55	68	63.8
2	98	107	77	96	94.5
Mean	60.7	77.7	54.3	68.0	

Ss were informed of the reduction in size of population from which the stimuli were selected. These results also supply a partial answer to Question 2 by indicating that set in the sense of a restriction of response may be produced in two different ways.

In view of the results in Experiment I, perhaps the most surprising outcome of the present experiment was the finding that temporal arrangement for presenting alternatives influenced identification. Further analysis indicated that the effect was limited to a particular set of conditions. Thus, only when two alternatives were employed was the viewing of the alternatives before and after the stimulus significantly superior to after ($F = 12.7$, $p < 0.01$). This result answers Question 6 (a) and supplies evidence in support of the hypothesis that set can augment the discrimination of stimulus-elements.

[10] G. A. Miller, G. A. Heise, and William Lichten, The intelligibility of speech as a function of the context of the test materials, *J. exp. Psychol.*, 41, 1951, 329-335.

[11] Krulee, Podell, and Ronco, *op. cit.*, 1954, 79.

Experiment III

Because of the unexpected results of Experiment II, Experiment III was deemed necessary. The specific aims of this experiment were (1) to determine whether the superiority of set given both before and after the stimulus over that given only after was a repeatable finding; and (2) to determine whether or not the superiority of set before and after presentation of the stimulus would be found in a situation which more clearly entailed the discrimination of stimulus-elements.

Method. Two different experimental situations were used. The first was identical to that of the previous experiments. Each *S* viewed stimuli consisting of single degraded letters projected on a screen. He was aided in the identification of these stimuli by being given sheets of paper on which two letters were printed. These sheets of paper were presented in the usual manner, *i.e.* either before and after or only after each stimulus. Final identification was indicated by circling the alternative that he believed to correspond to the stimulus.

The second experimental situation was somewhat different. Here each *S* viewed pairs of degraded letters projected on a screen. In addition he was informed as to the identity of one of the letters either before and after or only after the stimulus, that is, he was allowed to look at a single printed letter which corresponded to one of the two projected. His task was to determine which of those two letters corresponded to the letter printed on his sheet.

In the first experimental situation each of the 13 degraded letters used as stimuli was presented six times, making a total of 78 presentations. Groups of alternatives were prepared by pairing each one of the 13 letters with every other one, making a total of 78 pairs of alternatives. The pairs of alternatives were so combined with the stimuli that no alternative or spatial location on the sheets was correct more than any other.

In the case of the second experimental situation, every degraded letter was paired with every other one, again making a total of 78 presentations. Each of the 13 printed letters was so combined with the pairs of degraded letters that no particular letter or spatial position in the stimulus-field was correct more than any other.

In both experimental situations, stimuli of two levels of ambiguity were employed; these had been identified previously 16.5% and 3.5% of the time. Half of the *Ss* in each experimental situation studied their printed sheets both before and after the stimulus, and half only after. Other procedural characteristics were identical to those in Experiments I and II. Seven *Ss* were randomly assigned to each of the eight different combinations of experimental conditions, making a total of 56 *Ss*.

Results. The total number of identifications and the appropriate experimental conditions are presented in Table III. An analysis of variance performed on these data indicated that they were significantly influenced by all three variables — type of experimental situation ($F = 6.9$, $p < 0.05$), level of ambiguity ($F = 269.2$, $p < 0.01$) and temporal arrangement for giving the set ($F = 62.6$, $p < 0.01$).

That the presence of set before and after the stimulus was again su-

perior to its presence only after indicates that the earlier outcome of Experiment II is repeatable and was not dependent upon a particular selection of stimuli and *S*s. Moreover, because essentially the same results were found in the second experimental situation, the effect seems completely explicable in terms of the hypothesis offered earlier, *i.e.* an augmentation of the discrimination of important elements contained in the distorted stimuli.

In Experiment I the effects of set were found to be independent of the level of ambiguity. Such was not the case, however, in the present experi-

TABLE III

Total Number of Correct Identifications in Experiment III

(Each entry represents the total number of identifications made by seven *S*s on 78 stimuli.)

Situation	Level of ambiguity I: before and after	I: after	II: before and after	II: after
1	438	366	309	282
2	428	348	282	266

ment. The most reasonable explanation for this reversal seems to be the fact that a more restrictive set was employed in Experiment III.

The finding that more identifications were made in the first experimental situation than in the second cannot be interpreted. No prior prediction of this outcome was made because there was no basis for it.

Summary

The present experiments were designed to determine: (1) whether set varied in degree; (2) if its effects were related to level of ambiguity of the stimuli used; and (3) under what conditions set acts to increase perceptual accuracy by augmenting discrimination, by increasing interpretability, or by increasing correct responses simply by reducing the number of alternative responses.

In three experiments letter-patterns of varying levels of ambiguity were identified. *S*s were aided in this task by being differentially set. This was accomplished by providing a limited set of alternative responses varied in the following ways: (1) different numbers of alternatives; (2) alternatives drawn from different populations; and (3) alternatives in different temporal relation to the stimulus. These experimental operations yielded the following general results.

(1) Set varied in degree rather than being all-or-none in nature. This was shown by the fact that a reduction in number of alternatives or the substitution of alternatives from a restricted rather than unrestricted population increased the frequency of correct identification.

(2) The effect of set was influenced by the level of ambiguity if highly restrictive sets were used.

(3) In augmenting the identification of ambiguous stimuli, set was found to play three roles: (a) Set increased perceptual accuracy by augmenting the discrimination of potentially relevant elements or dimensions contained in the stimuli. Evidence for this role of set is the fact that under certain conditions a set both before and after the stimulus produced more identifications than one only after. This effect appeared only when a small number of alternatives were employed and stimulus-patterns were used containing discriminable elements. (b) Set increased the frequency of identification by reducing the number of possible responses. This was not the only role of set, however, because even when corrected for increases in chance expectancy, increases in identification were obtained with increases in degree of set. (c) Set, thus, also aided identification by making the residual elements in the degraded stimuli more interpretable. This was accomplished by supplying otherwise unavailable responses and by eliminating possibly competing ones.

AN EXPERIMENTAL ANALYSIS OF SET: THE ROLE OF SENSE-MODALITY

By EUGENE R. LONG, University of North Carolina, RICHARD H. HENNEMAN, University of Virginia, and WILLIAM D. GARVEY, Naval Research Laboratory

In an earlier study we reported the ability of a 'set' to alter the accuracy of reporting about a degraded stimulus-object. It was shown that set could be varied in degree; that its influence depends on the level of ambiguity of the degraded stimulus; and that it may affect discrimination, or interpretation, or it may simply make one response more probable by decreasing the alternatives.[1]

A number of questions raised in the postulate remain unanswered. The following experiments were designed to test whether a quite different method of establishing the set, namely, by ear rather than on a printed page, would change the results. The experiments were also designed to test words as stimulus-objects in place of letters.

There are evident differences in the functioning of the visual and auditory senses.[2] Because of such differences, set may function differently in the two modalities. A similar comparison can be made of words and letters. For example, words come from a larger population of stimuli than do letters of the alphabet. Thus, there is more inherent uncertainty in words than in letters. On the other hand, it might be argued that words possess more meaning than do letters.

EXPERIMENT I

The substitution of auditory stimuli for visual raises an additional question about what is the relationship between the modality of the stimulus and that of the set? In our previous experiments both were presented

* Received for publication June 9, 1959. This study is based upon a modification of a Technical Report written under Contract W 33 (038)-ac-21269 between the University of Virginia and the USAF, Psychological Branch, Aero-Medical Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio.

[1] E. R. Long, L. S. Reid, and R. H. Henneman, An experimental analysis of set: Variables influencing the identification of ambiguous, visual stimulus-objects, this JOURNAL, 73, 1960, 553-562.

[2] R. H. Henneman and E. R. Long, A comparison of the visual and auditory senses as channels for data presentation, *USAF WADC Tech. Rep.*, 53-363, 1954, 1-38; W. D. Garvey, The intelligibility of speeded speech, *J. exp. Psychol.*, 45, 1953, 102-106.

visually. If audition were used for both in the present experiment, any differences obtained in the functioning of set could well be due to either one or both of the two changes. To answer this question, distorted stimulus-words were presented by ear and the alternatives that furnished the set were presented both by ear and in printed form.

Method. The stimuli were 33 distorted spondaic (two-syllable, equally accented) nouns which were aurally presented. These words were distorted by recording them on a Brush Sound Mirror (Model B K 403) and then playing them back over the Sound Mirror with a 2.5 speed-up. This speed-up was accomplished by substituting a larger capstan for the smaller one originally employed for recording. Garvey previously reported that at a level of 42 db. above threshold this procedure produced intelligibility scores of approximately 8.0%. When the present investigators made a later check of this, however, by presenting these words to 10 additional *Ss*, only one correct identification out of a possible 330 was made, thus indicating an even lower intelligibility level for the *Ss* used in this experiment.

TABLE I

Sense-Modality and Temporal Arrangement for Presenting Alternatives

Sequence	Alternatives	
	Before the stimulus	After the stimulus
1	visual	visual
2	auditory	visual
3	none	visual
4	visual	auditory
5	auditory	auditory
6	none	auditory

Set was manipulated by presenting alternatives both aurally and visually. Aural presentation entailed recording groups of four spondaic words on tape, and then playing these back in undistorted form before, after, or both before and after, the presentation of the corresponding distorted stimulus-word. When the alternatives were presented visually, *Ss* viewed them on sheets of paper before, after, and both before and after, the stimulus-objects. The six sequences used for presenting alternatives are presented in Table I.

The procedure was so arranged that each *S* received one of the six sequences of alternatives that produced the set. During the first 7 sec. of a particular trial, the *Ss*, receiving alternatives before the stimulus, were allowed either to look at a sheet on which were printed four alternatives or to listen to the same four alternatives played back through earphones. Those *Ss* receiving no alternatives or receiving alternatives by ear looked at blank sheets of paper at this time. Then all *Ss* turned their sheets over. The stimulus was presented and immediately afterward those *Ss* who had visual alternatives after the stimulus looked at their appropriate sheets. Those receiving the auditory alternatives heard the four words played back. At the end of this time all the *Ss* then wrote down what they thought the stimulus-word had been. This procedure was repeated 33 times (once for each stimulus), requiring a total time of approximately 16 min.

Ten *Ss* were randomly assigned to each of the six different sequences of experimental conditions, making a total of 60 *Ss*.

Results. Frequencies of identification under each of the six conditions are presented in Table II. An analysis of variance performed on these data indicated that the presentation of alternatives both before and after the stimulus produced significantly more identifications than did the presentation only after ($F = 28.8$; $p < 0.01$). The modality used to present the alternatives before the stimulus, however, had no significant effect. The modality used for presenting the alternatives after the stimulus produced somewhat more complicated results. Vision was significantly superior to audition ($F = 24.3$; $p < 0.01$) if the alternatives had also been presented prior to the stimulus; if they had not been presented prior to the stimulus, vision and audition were equal.

TABLE II

TOTAL NUMBER OF CORRECT IDENTIFICATIONS IN EXPERIMENT I

(Each entry represents the total number of identifications made by 10 *Ss*. The total possible equals 330.)

Presentation after stimulus	Presentation before stimulus		
	Visual	Auditory	None
Visual	290	282	201
Auditory	233	219	198

These data clearly indicate that the identification of distorted words presented aurally can be improved by set. Moreover, this improvement also depends on the modality employed after the stimulus. The surprising outcome, in view of the previous results, is the finding that presenting the alternatives both before and after the stimulus is superior to presenting them only after. In the previous experiment this was shown only with fewer than four alternatives.

It could not be concluded that these results were solely the result of using auditory rather than visual stimuli. Words had been substituted for letters, as well as auditory for visual stimuli. So another experiment was conducted. The specific purpose of this experiment was to assess the relative contribution of using words instead of letters and of different durations of stimuli—⅓ sec. instead of 4.0 sec.

EXPERIMENT II

In the next experiment, the same stimulus-objects (spondaic words) were employed, but they were presented visually in distorted form. This was

done by first photographing them out of focus and then projecting the resulting 35 mm. transparencies with the projector misfocused. Two durations of exposure were employed : 4.0 sec. and ⅓ sec. Set was manipulated by the use of the same groups of alternative responses. These were presented both before and after each stimulus or only after, and they were always presented visually.

Results. Frequencies of identification under each of the conditions are presented in Table III. An analysis of variance indicated that set significantly increased the number of identifications ($F = 88.5$, $p < 0.01$). Presenting the alternatives both before and after the stimulus, however, was not significantly better than presenting them only after. Similarly, duration of stimuli had no significant effect. These results indicate that the identification of distorted *words*, presented either aurally or visually, may be facilitated by giving an appropriate set. Words are comparable with letters. They also indicate that the superiority of presenting aural alternatives both before and after, over that of only after, was indeed due to the use of auditory stimuli rather than the use of words or short durations of presentation.

TABLE III

TOTAL NUMBER OF CORRECT IDENTIFICATIONS IN EXPERIMENT II

(Each entry represents the total number of identifications made by 5 *S*s. Possible total equals 165.)

Duration of exposure	Alternatives		
	both before and after stimulus	after stimulus	None
0.33 sec.	111	109	42
4.0 sec.	109	98	35

SUMMARY AND CONCLUSIONS

In the present experiments, the *S*s were required to identify distorted spondaic words, which were presented through earphones. Set, designed to reduce ambiguity, was manipulated by presenting groups of four words, one of which corresponded to the distorted stimulus-word. These procedures yielded the following results.

(1) Presentation of alternatives both before and after the stimulus produced more identifications than presentation only after.

(2) No single relationship was found to hold between the modality used to produce the set and that used as the stimulus. Vision and audition were equally effective when used before the stimulus. After the stimulus, vision was superior to audition if the alternatives had also been presented before

the stimulus. If they had not been presented before the stimulus, audition was as good as vision.

(3) When the same spondaic words were presented visually, the advantage of alternatives both before and after the stimulus disappeared. This suggests that the superiority found with aural material was contingent upon the use of auditory stimuli rather than on the use of words or the use of short durations.

(4) In augmenting the identification of distorted auditory stimuli, set was again seen as playing three roles: (a) it increased the probability of each response alternative simply by reducing the number of alternatives; (b) it increased the interpretability of residual stimulus-elements; and (c) it aided the discrimination of important stimulus-elements. The latter role was demonstrated by the fact that aural alternatives both before and after the stimulus were more effective than alternatives—only after the stimulus.

AN EXPERIMENTAL ANALYSIS OF SET: THE EFFECT OF CATEGORICAL RESTRICTION

By L. STARLING REID, RICHARD H. HENNEMAN, University of Virginia, and EUGENE R. LONG, University of North Carolina

In experiments previously reported, 'set' was manipulated by employing different numbers of alternative responses; by selecting the alternatives from different populations; and by presenting the alternatives in different temporal arrangements.[1] The purpose of the present study is to determine whether set can be effectively increased by greater categorical restriction, *i.e.* by giving the *S*s more attributes of the category to which a stimulus-word belongs.

The following questions were posed: (1) Does the restriction of the specific category to which a degraded stimulus-word belongs increase the frequency of its correct identification?

(2) If categorical restriction does increase the frequency of identification, is this increase contingent upon the use of the categorizing terms both before and after the stimulus or is it equally effective if presented only after the stimulus?

(3) Is familiarity with the specific words used as stimuli, apart from knowledge of the categories, a factor in the effectiveness of this method?

METHOD

Forty-eight words, all generally related to sports, were used as stimuli. These words were first typed on individual cards. They were then made equal in length by typing x's before the first letter and after the last letter of each word, as shown in the following examples: X D O D G E R S X and X X X C O B B X X. The cards were then photographed with the camera misfocused. This procedure resulted in the production of 35-mm. transparencies which, when projected out of focus, were correctly identified only 2% of the time without the aid of set.

The 48 stimulus-words were selected so that they could be subdivided into successively smaller categories. Thus, all the words were related to sports, one-half were related to baseball and the other half to football. Each of these categories was

* Received for publication June 9, 1959. This study is based upon a Technical Report written under Contract W 33 (038)-ac-21269 between the University of Virginia and the USAF, Psychology Branch, Aero-Medical Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio.

[1] E. R. Long, L. S. Reid, and R. H. Henneman, An experimental analysis of set: Variables influencing the identification of ambiguous visual stimuli, this JOURNAL, 73, 1960, 553-562; E. R. Long, R. H. Henneman, and W. D. Garvey, An experimental analysis of set: The role of sense modality, this JOURNAL, 73, 1960, 563-567.

further subdivided into names and terms. Names were broken down into teams and players, and terms into positions and plays. Tables I and II list all of the subcategories and the stimulus-words in each category.

Prior to the *Ss'* attempt to identify the words, two types of familiarization were provided. One group of *Ss* was allowed to study both the categories and the stimulus-words contained in each. The second group was allowed to study only the categories. Both groups were given a 5-min. study-period, then a written test of

TABLE I

ORGANIZATION OF THE FIRST PRINCIPAL CATEGORY UNDER THE GENERAL CLASS, SPORTS

Baseball							
Names				Terms			
Teams		Players		Positions		Plays	
National	American	Past	Present	Infield	Outfield	Offensive	Defensive
Dodgers	Yanks	Ruth	Shantz	First	Left	Hit	Catch
Pirates	Tigers	Cobb	Berra	Second	Center	Walk	Throw
Giants	Indians	Speaker	Snider	Third	Right	Steal	Field

TABLE II

ORGANIZATION OF THE SECOND PRINCIPAL CATEGORY UNDER THE GENERAL CLASS, SPORTS

Football							
Names				Terms			
Teams		Players		Positions		Plays	
Pro	College	Past	Present	Lineman	Backfield	Offensive	Defensive
Rams	Irish	Grange	Walker	Guard	Quarter	Block	Rush
Bears	Trojans	Luckman	Baugh	Tackle	Half	Pass	Recover
Lions	Sooners	Thorpe	Graham	End	Full	Plunge	Punt

recall, a second 5-min. study-period, at the end of which the experiment proper was begun.

Degree of set was varied by giving *S* a number of categorizing terms ranging from one to five. Thus, those *Ss* receiving the minimal set saw only one word, *sports*, printed on their sheets. Those receiving the next higher degree of set saw two words, *sports-baseball* or *sports-football*. Those *Ss* having the highest degree of set saw five successive descriptive or categorizing words, thus *sports-football-name-team-pro*, the specific combination of terms again being determined by the particular stimulus-word being projected.

The experimental procedure was like that of the previous experiments and was so arranged that the categorizing words might be viewed both before and after or only after the stimulus. This procedure was repeated 48 times, once for each stimulus-word, and required a total time of approximately 25 min.

Sixty *Ss* (college men) were randomly assigned to the 20 different combinations of experimental conditions.

Results

The total numbers of correct identifications are presented in Table III. An analysis of variance performed on these data indicated that (1) greater restriction in the category resulted in a significant increase in identification ($F = 9.7$, $p < 0.01$); (2) familiarity with both the stimulus-words and categories produced significantly more identifications than did familiarity with the categories alone ($F = 13.4$, $p < 0.01$); and (3) the temporal arrangement for presenting the categorizing words did not significantly influence identification.

TABLE III

Total Number of Correct Identifications

(Each entry represents the total of three *S*s.)

Type of familiarization	Degree of set: I		II		III		IV		V	
	words plus categories	categories only	words plus categories	categories only	words plus categories	categories only	words plus categories	categories only	words plus categories	categories only
Set after stimulus-word	59.1	45.0	48.9	39.9	50.1	48.9	90.0	78.0	93	36.9
Set before and after stimulus-word	57.0	56.1	50.1	36.0	72.0	51.0	111.0	72.9	96	75.0

These facts indicate that set can be effectively varied by manipulating the size of a category and need not entail varying the specific alternative responses allowed *S*. Nevertheless, the categorical restriction does reduce the number of appropriate stimulus-words, and thus the present results are in accord with those of Miller, Heise, and Lichten, which indicate that success of identification is a function of the size of the population of the stimulus-objects.[2]

The effect of categorical restriction was independent of when the categorizing terms were presented. In an earlier article the present investigators reported that restricting responses both before and after the presentation of the stimulus-objects was no better than that only after, when more than two alternatives were employed.[3] It may be, therefore, that the corresponding finding in the present experiment is also due to the use of as many as four alternative responses.

As an alternative, it might be hypothesized that the lack of any influence

[2] G. A. Miller, G. A. Heise, and William Lichten, The intelligibility of speech as a function of the context of the test materials, *J. exp. Psychol.*, 41, 1951, 329-335.

[3] Long, Reid, and Henneman, *supra*, 553-562.

of a pre-exposure set is a result of the particular way in which the stimuli were degraded, *i.e.* misfocusing. This hypothesis is supported by the data of Lawrence and Coles and by those of a second earlier experiment in this present series.[4] It is possible that misfocusing degrades the entire stimulus. If the stimuli had been degraded by partial omission or by the simultaneous addition of other stimuli, the result would have been that some of the elements of the projected stimuli were intact. In such a case it would still be possible to discriminate some of the residual elements of the stimulus, and set could perhaps augment this. On the other hand, when misfocusing is employed and all elements have been degraded, there seems to be less chance for augmenting this discrimination.

The fact that the group previously acquainted with both words and categories is superior to the group which studied only the categories indicates that previous familiarity with the stimuli themselves aids in their identification. This is certainly not unexpected, being in keeping with the findings of others who have used other types of stimuli and other types of impoverishment.[5] A somewhat surprising outcome of the present experiment, however, is the fact that degree of categorical restriction and type of previous familiarization did not interact. Thus, increases in categorical restriction did not improve the identification of the *S*s previously familiarized with categories and stimulus-words more than those familiarized only with the categories.

SUMMARY

The purpose of the present experiment was to determine whether, in the identification of ambiguous stimuli, could be effective by restricting the category of the response. The principal findings were the following:

(1) An increase in categorical restriction was accompanied by an increase in frequency of correct identification.

---

[4] D. H. Lawrence and G. R. Coles, Accuracy of recognition with alternatives before and after the stimulus, *J. exp. Psychol.*, 47, 1954, 208-214; Long, Henneman, and Garvey, *supra*, 563-567.

[5] No complete review of the extensive literature on the relationship between degree of learning and accuracy of perceptual recognition will be attempted here. The following are representative: D. H. Howes and R. L. Solomon, Visual duration threshold as a function of word probability, *J. exp. Psychol.*, 41, 1951, 401-410; R. L. Solomon and L. J. Postman, Frequency of usage as a determinant of recognition thresholds for words, *ibid.*, 43, 1952, 195-201; L. J. Postman and M. R. Rosenzweig, Practice and transfer in visual and auditory recognition of verbal stimuli, this JOURNAL, 69, 1956, 209-226; Rosenzweig and Postman, Frequency of usage and the perception of words, *Science*, 127, 1958, 263-266; Israel Goldiamond and W. F. Hawkins, Vexierversuch: Word-frequency and recognition, *J. exp. Psychol.*, 56, 1958, 457-463.

(2) Presenting the categorizing terms before stimulation neither significantly improved identification, nor interacted in its effect with degree of categorical restriction.

(3) The *S*s who had been familiarized with both the stimulus-words and categories were significantly superior to those who had studied only the categories. Familiarity with the specific stimulus-words was not, however, a factor in the first finding above. The groups with both levels of familiarity demonstrated essentially the same general increase in correct identification as the degree of categorical restriction was increased.

THE STABILITY OF A STANDARD OF LOUDNESS AS MEASURED BY COMPENSATORY TRACKING

By J. DONALD HARRIS and ANDREW G. PIKLER,
U. S. Naval Medical Research Laboratory
New London, Connecticut

When an *S* listens to a sound changing in loudness, any attempt to manipulate the system so that programmed changes are minimized can be thought of as compensatory tracking. A form of this task, actually a limiting case in which zero loudness is maintained as the auditory stimulus sweeps from low to high frequencies, has been standardized in Békésy's audiometry.

Zwicker and Feldtkeller used a form of tracking to produce isophonic contours. *S* was asked to maintain a constant balance in loudness between a pure tone of 1000 ~, constant in intensity, and a tone changing automatically in frequency from low to high, the loudness of which was under *S*'s control.[1]

Pikler pointed out that compensatory tasks and pursuit-tracking could be undertaken at supra-threshold levels solely on the basis of subjective standards.[2] For example, *S* could be required to maintain constant loudness in the face of intensive changes programmed in a tone of constant frequency.

The present experiment was designed to explore the stability of 1000 ~ at 40-phon loudness for short periods of compensatory tracking.

APPARATUS AND PROCEDURE

A Békésy audiometer, with a continuously variable pen-speed, was used to generate five programs. These programs consisted of semi-random intensive changes introduced into a tone of 1000 ~ at four different attenuated speeds (0.35, 0.58, 0.76, and 1 db. per sec.) which were encompassed within a 20-db. range. All these programs were preceded for the first 3 min. by a steady 1000 ~ signal which permitted *S* to reach maximal, prestimulatory adaptation. A momentary break in the tone, toward the end of the 3-min. period, warned *S* that he could expect a change in loudness within 8–12 sec.

The programs were recorded on an Ampex Model 300 studio tape-recorder. They were accurate, as demonstrated by a sound-power-level-recorder, to within 0.5% in duration and 0.5 db. in intensity of the values given in Table I.

During an experimental session, a program was presented monaurally to *S* by

* Received for publication August 20, 1959.

[1] E. Zwicker and R. Feldtkeller, Ueber die Lautstärke von gleichformigen Geräuschen, *Acustica*, 5, 1955, 303-316.

[2] A. G. Pikler, Patent Disclosure No. 21, 197 to U. S. Navy Bureau of Ships, October, 1955.

means of a decibel step-attenuator and a rotary 100-db. attenuator geared to a change of 1 db. for each 15° rotation. A very light, smoothly working load, with a negligible backlash, was presented to *S* by the rotary attenuator as he manipulated a wheel 2 in. in diameter. A clockwise rotation increased the intensity at the earphone (Permoflux Co. PDR-8), and a counterclockwise rotation decreased it. *S* could not obtain visual or other cues concerning the dial-position.

S's task was to maintain a constant level of loudness by adjusting the rotary-dial. Prior to the experimental sessions, *S* was given adequate practice to reach his asymptote of audiomotor skill.

TABLE I

CONSTRUCTION OF PROGRAMS 1–5

Program (in sec.)

1 (0.35 db./sec.)		2 (0.58 db./sec.		3 (0.76 db./sec.)		4 (1 db./sec.)		5 (1 db./sec.)	
flat	10	flat	10	flat	10	flat	10	flat	10
crescendo	20	cresc.	8	decr.	7	cresc.	6	cresc.	11.6
decrescendo	38	decr.	12	cresc.	13.6	decr.	5.3	decr.	10
cresc.	16.6	cresc.	16	decr.	9.6	flat	8.6	cresc.	8.3
flat	9	decr.	19	flat	13	decr.	6	flat	10.6
cresc.	21	cresc.	24	cresc.	14.5	cresc.	8.6	cresc.	16.6
decr.	40	decr.	18.6	decr.	11	decr.	5	decr.	12.5
cresc.	13.3	flat	8.6	flat	8.6	flat	9	cresc.	10.6
		cresc.	7.6	decr.	7	cresc.	8	flat	14.5
Total time	168 sec.			cresc.	18.3	decr.	15.6	decr.	10
			124 sec.	decr.	14.6	cresc.	10	cresc.	12.5
				flat	8	flat	9	flat	11
					135.5 sec.		101.3 sec.		138.5 sec.

To record *S*'s response, the rotations of his wheel were transferred to a voltage-recorder. To accomplish this a second matched rotary 100-db. attenuator was arranged on the same shaft. Through this second attenuator was led a convenient voltage from an independent source of the voltage-recorder. This recorder had a paper speed of 9-in. per sec. with a 40-db. dynamic range through a 4-in. excursion. Thus a complete visual record of *S*'s dial-adjustments could be obtained.

In a control experiment, *S* was provided monaurally or interaurally with a tone of constant-intensity and different frequency to serve as a loudness-reference. After a test-run, the program-tape and the paper in the voltage-recorder were rewound to the 'start' positions and the tape-recorder led directly to the voltage-recorder. In this way the program itself could be superimposed in different-colored ink directly over *S*'s response for fine comparisons. (It was only necessary that the recording-pen for the initial phase of the program be adjusted to the same baseline as had been used for *S*'s response.) The error of asynchrony between the program and *S*'s response was established repeatedly to be less than 0.5 sec.

The procedure for quantifying *S*'s errors of response was simple. Fig. 1 shows a typical record of a test-run. At any moment the error in decibels could be measured in vertical distance between the program and the compensatory response. The area between the two tracings is measured to obtain the cumulated error. The method of

doing this was to cut out the relevant area for each run and to weight it to the nearest milligram. As a reference for the unit of cumulated error, the weight of an area representing 1 db. × 1 sec. was repeatedly measured. The total weight of the cumulated error divided by the weight of the unit area gave the error in terms of the unit (db.-sec.) for which the word 'mistrack' was coined. This procedure neglects sign, but a preliminary visual assessment could be had of any tendency to overshoot or undershoot the program. Finally, the total 'mistracks' for any test-run could be divided by the number of seconds in that program to yield the averaged

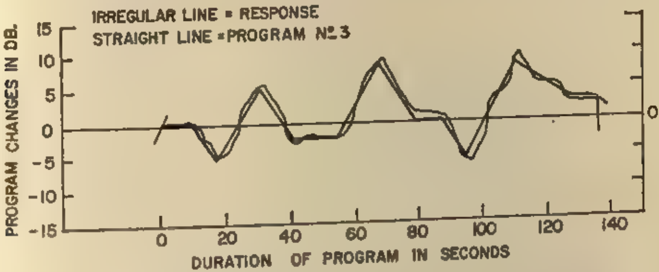


Fig. 1. Program 3, with Typical Response Superimposed

momentary error in decibels. Time-lags in response were sampled at the peaks and points of transition.

Results

The averaged momentary error, as related to speed of attenuation for 5 *Ss* with normal hearing at 1000~, is shown in Fig. 2. Two *Ss* (*JDH* and *CRP*) show a tendency toward better performance with increased speed of attenuation up to a limit of 0.76 db. per sec., which would appear to represent their personal asymptote. Two *Ss* (*RHE* and *AGP*) do not, however, confirm this trend.

Table II shows the averaged momentary errors of all *Ss* for all runs. It is seen from Columns 1–4 that the average errors for the four programs are indistinguishable (means for all 15 runs per program of 1.9–2.0 db.) We may, then lump the programs in considering the accuracy and consistency of this type of tracking among our *Ss*. Column 5 shows individual differences among the *Ss* in their ability to track (means of 1.1–2.8 db.), and Column 6, their consistency (*SD* 0.52–1.80 db.).

These data represent relatively very efficient performances, particularly, when we take into consideration the fact that the data were obtained at 1000~ at 40 phons, and that in the field of loudness no analogue of absolute pitch exists.

To ascertain the level of performance attainable by untrained *Ss*, a dozen young men of average intelligence underwent a session with Program 2. All understood the instructions and performed creditably. A distribution of their momentary errors averaged as follows: 3.2, 3.5, 3.6, 3.9, 4.0, 4.2, 4.2, 4.4, 5.4, 5.5, 6.2, 7.3 db.

When four of these men were accorded a second trial, two did not improve, but one improved from 5.4 to 2.4 db., and another from 4.2 to 1.6 db. As we knew from preliminary training-sessions with our *Ss*, several trials are sometimes needed

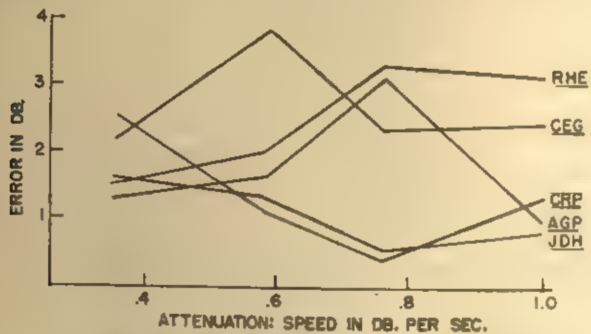


FIG. 2. AVERAGED MOMENTARY ERRORS OF TRACKING IN *DB.* AS A FUNCTION OF SPEED OF ATTENUATION

TABLE II

AVERAGED MOMENTARY ERROR IN DECIBELS

Ss	Program (three consecutive runs) 1			2			3			4			Mean	*SD*	Program 3 (control)		
JDH	1.1	2.0	1.8	1.3	0.8	1.9	0.8	0.4	0.7	0.7	1.2	0.6	1.1	0.52	0.9	1.8	3.2
RHE	1.3	2.0	1.5	1.3	2.0	2.7	7.1	1.8	1.0	2.6	1.3	5.5	2.5	1.80	2.6	1.3	5.5
CRP	1.2	3.5	2.7	1.5	1.0	0.6	0.3	1.0	0.7	1.2	0.8	1.9	1.4	0.89	—	—	—
CEG	2.6	2.6	2.8	3.5	2.6	5.5	2.5	1.5	3.0	3.6	2.4	1.3	2.8	1.03	2.7	2.6	3.3
AGP	0.8	1.4	1.6	1.7	2.0	1.1	2.2	1.5	5.7	1.1	4.0	0.9	2.0	1.38	1.3	2.0	1.3
Mean		1.9			2.0			2.0			1.9		1.96				

to get the 'feel' of the task; and even with practice an *S* can, from time to time, get rather far off the standard.

The time-lags of the *Ss'* responses, as a function of speed of attenuation, are shown in Fig. 3. Four *Ss* showed a marked tendency to decrease the time-lag up to a limit of 0.76 db. per sec. Beyond this point, only one *S* continued the same trend. The mean time-lag shows a decrease up to 0.76 db. per sec.

The time-lags may be expected to vary not only with speed of attenuation but also with the transitions from plateau to crescendo, decrescendo to plateau, crescendo to decrescendo, or their opposites. The data were analyzed with this possibility in mind, but within the limits of the parameters and precision of measurement no clearcut tendency appears for

one type of transition to yield a different average time-lag than any other type.

Another look at the stability of the standard of subjective loudness is provided by a comparison with the tracking-error yielded when the same or the opposite ear is furnished with a control tone of constant loudness and of different frequency against which continuous matches of loudness may be made. The pertinent data are shown in the extreme right column

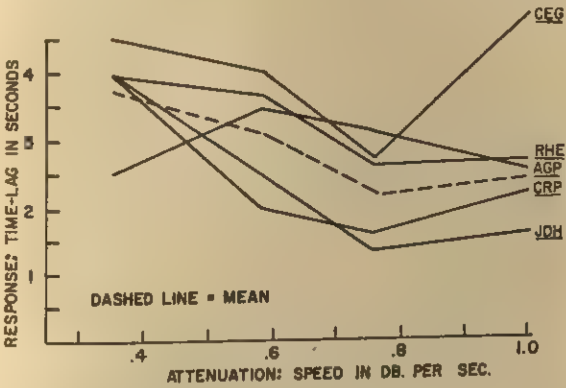


FIG. 3. TIME-LAG OF RESPONSES IN SEC. AS A FUNCTION OF SPEED OF ATTENUATION

of Table II. No improvement is seen for 4 *S*s when the reference-tone is provided.

Still another look at the stability of the subjective loudness-standard is provided by *S*'s resistance to programmed drifts which do not hover symmetrically around the original loudness but progress systematically in semi-random fashion toward either softer or louder average levels. In such a case, *S* can build up a picture of the program as being asymmetrical since he can easily sense that he is making more clockwise or counter-clockwise rotations as the case may be. In these trials we are primarily interested in how far off, if at all, *S* can be drawn from his internal loudness-reference by some cumulation of systematic tracking-errors. Note that Program 5 contains three plateaus. The averaged momentary errors over these plateaus, during portions which neglect the error due to time-lag, can be compared with the errors over the three plateaus of Program 3. These two programs were constructed with a view to making just this comparison. Any real increase or decrease of error for Program 5 as against Program 3 could be construed as indicating an effect of the asymmetry of

the program. Table III compares the momentary errors for the three plateaus of Programs 3 and 5 for the five *S*s. *AGP*, *CEG*, and *RHE* show larger mean errors on Program 3, while *CRP* and *JDH* show slight errors on Program 5. Only one *S* (*RHE*) seems to show a significant difference at the 5% level, because of his anomalous performances on the first plateau of Program 3. The authors conclude that asymmetry has no appreciable effect within a 20-db. dynamic range.

It is established that the practice during the experimental sessions had

TABLE III

Averaged Momentary Errors in Db. for a Symmetric and an Asymmetric Program

	Program 3 (symmetric)			Program 5 (asymmetric)			
S	1st	2nd plat.	3rd	1st	2nd plat.	3rd	*t*
AGP	1.2	4.2	7.3	0.8	0.2	1.3	1.98
	1.7	1.0	8.3	1.5	0.8	0.7	
	0	0	6.8	1.7	1.2	1.2	
CRP	0.5	2.0	0.7	0.7	1.0	2.3	0.49
	1.0	2.0	0.4	2.0	2.0	2.7	
	0	1.0	0.5	1.3	0.8	2.5	
JDH	0.6	0.5	0	0.7	0.5	0	0.34
	1.3	0.2	0	0.5	0.5	0.5	
	0.3	0	0.5	1.0	1.0	1.0	
CEG	2.5	3.0	3.0	1.5	1.5	1.0	0.36
	3.5	0	4.5	7.5	0.8	1.0	
	2.5	4.0	5.2	6.3	2.0	*	
RHE	6.2	3.2	3.0	2.3	2.0	3.3	2.57†
	8.0	2.5	0	2.2	1.0	0.7	
	9.0	3.5	2.5	0.8	1.5	0.8	

* Sample not interpretable. † Significant at the 5% level.

no effect. An inspection of Table II reveals that after combining the data of all the *S*s, the averaged momentary errors on the first, second, and third runs (replicas) were 1.9, 1.8, and 2.2 db.

Discussion

The present experiment might be thought to have some connection with servo systems in that *S*, while manipulating the rotary attenuator, also regulates the voltage from an oscillator. This second voltage could be used to close a loop to the source. Nothing of this sort, however, was done hence no servo mechanisms are involved.

The consistency of tracking compensatory loudness without objective reference is perhaps the most striking feature of these results. It is clearly shown that *S*s are attentive within a very short time to any type of loud-

ness-change and also that they compensate for such changes with accuracy and consistency. Valid data on the drift of subjective standards of loudness and their stability in the face of programmed drifts can be obtained without much if any previous training.

Of the two possible sources of variance contributing to the cumulative error; namely, auditory sensitivity and motor skill, it is our opinion that the present experiment, through training and by simplifying conditions, eliminated the variance due to motor skill. The results probably reflect the variance due to auditory sensitivity.

It is not especially fruitful to compare the size of *S*'s average momentary error with the size of his *DL* or *JND* for loudness. Tracking compensatory loudness is a psychophysical method itself in which memory of loudness and differential discrimination are somehow merged. Perhaps the closest analogy comes from the Method of Single Stimuli. With this method Pollack found a *DL* of about 0.6 db. even with a 10-sec. interval between tones (data obtained at 1000~ at 40-phon level).[3]

Still less are the data for the time-lag of the responses analogous in any close sense with classical reaction-time for auditory discrimination. *S* does not limit himself to a single movement but continuously initiates, inhibits, and reverses movements which must then be slightly accelerated until the effects of the time-lag are compensated.

A variety of uses in supra-threshold audiometry is encouraged by thsee data. The stability of the standards of subjective loudness of some of our *S*s is such that tracking compensatory loudness could serve to explore the inherent drift of the standard. Not only the stability of the loudness-standard, but pitch and many other aspects of the auditory stimulus could be handled. Even rather elaborate tasks might yield to this treatment, such as the extremely complex pursuit and compensatory trackings of pitch, loudness, and time simultaneously, as in the case of ensemble music.[4]

SUMMARY

Five *S*s listened to tape-recorded programs which presented variations in intensity of 0.3 to 1 db. per sec., within a 20-db. range, of a 1000~-tone at 40 phons. The *S*s attempted to achieve compensatory tracking, *i.e.* to maintain the loudness-constant by rotating a dual attenuator. The latter

[3] I. Pollack, Intensity discrimination thresholds under several psychophysical procedures, *J. acoust. Soc. Amer.*, 26, 1954, 1056-1059.

[4] A. G. Pikler, Auditory tracking and ensemble music, *Amer. Psychologist*, 13, 1958, 376; A. G. Pikler and J. D. Harris, Compensatory and pursuit tracking of loudness, *J. acoust. Soc. Amer.*, 32, 1960, 1129-1133.

controlled the loudness of the tone, and in addition allowed the tracking-responses to appear on the paper tape of a voltage-recorder. The programs themselves were also written on the paper.

The averaged momentary tracking-error was about 2 db. No difference in accuracy appeared among programs with different speeds of attenuation.

A simultaneous tone of constant intensity but of different frequency did not improve performance. Asymmetrical patterning of the program around the initial level did not affect the tracking-error. The average time-lag of the responses was about 3 sec.

In view of the accuracy and consistency of the data, suggestions were given for use of the method of auditory tracking in audiometry and engineering.

THE DISCRIMINABILITY OF TONES USED TO TEST STIMULUS-GENERALIZATION

By ALEC J. SLIVINSKE and JOHN F. HALL, Pennsylvania State University

The concept of stimulus-generalization depends upon the demonstration that the various stimuli can be discriminated in a situation other than the test of generalization. It has been tacitly assumed that the auditory stimuli, such as those used by Hovland in studies of generalization, although unquestionably discriminable in terms of differential judgments, would also be discriminable if tested by other psychophysical techniques.[1]

Investigations of Pollack and Garner, in which absolute judgments were utilized, resulted in a surprisingly small number of discriminable stimulus-categories over the auditory dimensions of pitch and loudness.[2] Both studies agreed in the finding that on the average only 4–6 discriminable categories were possible over ranges of 1000–8000 ~ at intensities of 5–100 db. Inasmuch as the test of generalization does not incorporate differential discrimination, estimates of stimulus-discriminability based on absolute judgments would appear to offer the more appropriate technique for determining the significance to be attached to the test of generalization. When the small ranges of values used in the studies by Hovland (150–2000 ~ at 40–86 db.) are contrasted to the larger ranges that Pollack and Garner found necessary for reliable absolute judgments, Hovland's choice of stimuli seems too narrow.

The present study was designed to investigate the discriminability of those stimuli utilized in Hovland's study of generalization, employing (a) the method of absolute judgments, and (b) a modified method designed to approximate the conditions of generalized tests. The results of the two

---

* Received for publication November 18, 1958. This study was supported by a grant from the National Science Foundation. The authors express their gratitude to Mr. Alvin Ugelow and Mrs. Lorraine Low for their assistance.

[1] C. I. Hovland, The generalization of conditioned rresponses: I. The sensory generalization of conditioned responses with varying frequencies of tone, *J. gen. Psychol.*, 17, 1937, 125-138; The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensities of tone, *J. genet. Psychol.*, 51, 1937, 279-291; R. A. Littman, Conditioned generalization of the galvanic skin response to tone, *J. exp. Psychol.*, 39, 1949, 868-882; D. D. Wickens, H. M. Schroder, and J. D. Snide, Primary stimulus generalization of the GSR under two conditions, *ibid.*, 47, 1954, 52-56.

[2] Irvin Pollack, The information of elementary auditory displays, *J. acoust. Soc. Amer.*, 24, 1952, 745-749; W. R. Garner, An informational analysis of absolute judgments of loudness, *J. exp. Psychol.*, 46, 1953, 373-380.

techniques were compared. The investigation was confined to the single auditory dimension of loudness.

The Pollack and Garner studies suggest that the ranges Hovland employed would not provide four absolutely discriminable categories. A third objective of the study was to provide a set of stimuli which might be maximally discriminable. To this end two modification were attempted. The first was to increase the range of stimuli employed, and the second was to select stimuli from a scale of equal discriminability based on absolute judgments over this wider range.

Method

Subjects. The *Ss* were 150 undergraduates from introductory courses in psychology. They were screened by a simple test to insure that they could hear a tone 26 db. SPL.

Apparatus. The apparatus consisted of an oscillator, PDR-8 earphones, and two timers to control the duration of the tones. The *Ss* were seated in a sound-shielded room in which the ambient noise was approximately 20 db. SPL. A 1000 ~ tone was used in all cases; its duration was always 0.6 sec. Transients in the onset and offset of tones were eliminated by the electronic switch of the oscillator.

Procedure. The general plan of the experiment was to apply the method of absolute judgments and the modified psychophysical method to the Hovland series of intensities and to a series arranged in equally discriminable steps over a much wider range. Data were first collected on the Hovland tones, *i.e.* 40, 60, 74, and 86 db. SPL. Seventy-five *Ss* were randomly assigned to five groups of 15. One group made absolute judgments and the other four groups used the modified method.

The group giving absolute judgments followed a procedure adapted from that described by Garner and Hake.[3] Each of the four intensities was presented 24 times in a random order. *Ss* assigned a number from one to four to the four tones from the softest to the loudest. The modified method differed in that one of the tones was designated as the standard stimulus. At the beginning of the experimental session the appropriate standard stimulus was presented three times to *S*. If *S* desired, it was presented again until he said he was familiar with it. These *Ss* were then told they would hear a series of tones differing in loudness and that they were to respond 'Same' or 'Different' to each tone depending upon whether they judged a given tone to be the standard or a different tone. The *Ss* were not told that there were only four tones in the series.

At the beginning of a session for the group giving absolute judgments, each of the four tones in the series was presented three times and given its appropriate number. Upon *S*'s request they were presented again until familiarity with them was established.

All 96 judgments were made in a 50-min. session in which a 2-min. rest-period was introduced halfway through the order. Before completing the second half of the judgments, all *Ss* were again familiarized with the standard stimulus or the whole series used for absolute judgments.

[3] W. R. Garner and H. W. Hake, The amount of information in absolute judgments, *Psychol. Rev.*, 58, 1951, 446-459.

The same procedure was replicated on four tones taken from an equal discriminability scale. The remaining 75 *Ss* were randomly assigned as before into five groups of 15 each. The equal discriminability scale was obtained from an earlier group of 20 *Ss*. The scale is shown in Fig. 1. The four intensities selected for the second series were 30, 66, 94, and 110 db.

Results

The data for *Ss* using the modified method consist of the frequency with which each tone was judged to be the same as the standard stimulus.

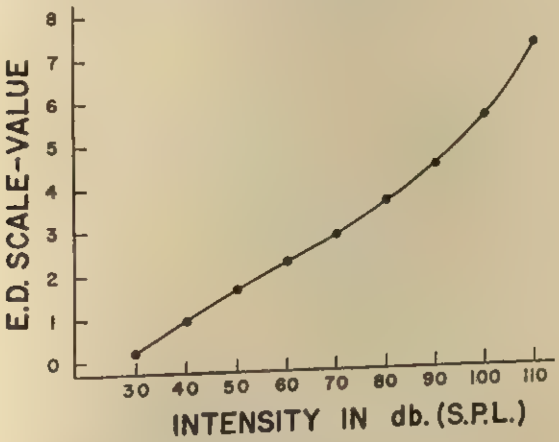


FIG. 1. AN EQUAL SCALE OF DISCRIMINABILITY FOR LOUDNESS

(Data were pooled for 20 *Ss*.)

The resulting curves showing means for the various groups are shown in Fig. 2. The four graphs show the distribution for the four standard stimuli. Along the abscissas are plotted the four variable stimulus-intensities. The ordinates show the mean percentage of 24 presentations to each of 15 *Ss* that each stimulus was judged the same as the standard stimulus. It is obvious that the intensity equal to the standard was correctly reported in a large number of instances. There is also little difference between the two ranges of stimuli.

The data of the absolute judgment groups were tabulated by summing the frequency of each response to each of the four stimuli for the two frequency-ranges. These data are provided in Fig. 3. Each graph shows the mean percentage with which each response category was used when one of the four stimulus-tones was presented.

Differences for both ranges were tested with analysis of variance of the number of response categories occurring to each of the four tones. Again a separate analysis was done for each of the two ranges, but in this case *Ss* and stimuli were not independent. Consequently, the variance due to *Ss*

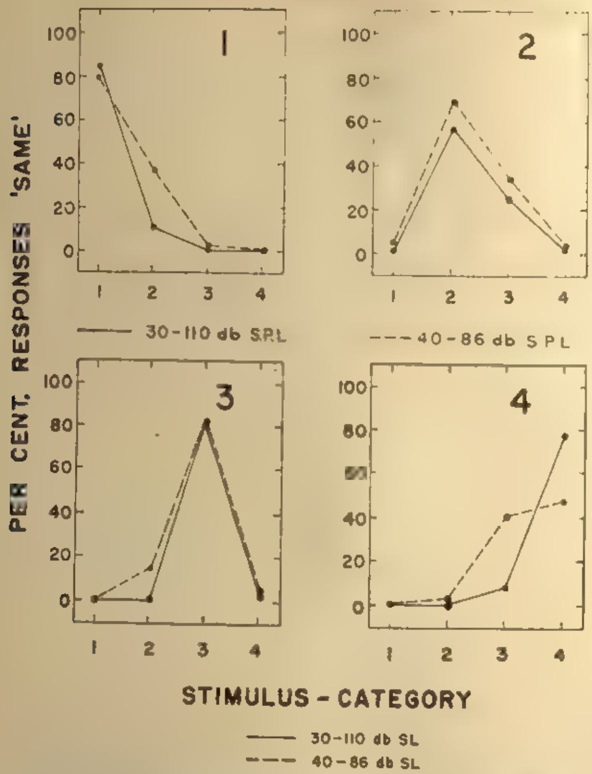


FIG. 2. THE PERCENTAGE WITH WHICH THE *Ss* JUDGED THE VARIOUS TONES TO BE THE SAME AS A STANDARD TONE

(The tone used as the standard is indicated by the larger number in each group.)

was tested in each analysis. The variance due to *Ss* was not significant in either range (*F*-ratio for Hovland range was 1.47 and for 30–110-db. range, 0.20 with 14 and 42 df.). The *F*-ratios for stimuli, as would be expected, are significant in both ranges. With 3 and 42 df. in each case, the ratio for the Hovland range was 5.28 ($p < 0.01$) and for the 30–100-db. range 3.68 ($p < 0.05$).

Discussion

A requirement of the stimuli used in an experiment testing stimulus-generalization is that they be discriminable in an absolute sense, that is, when presented in isolation. Our results clearly demonstrate that the Hovland tones are not completely discriminable in this sense. In fact, it may

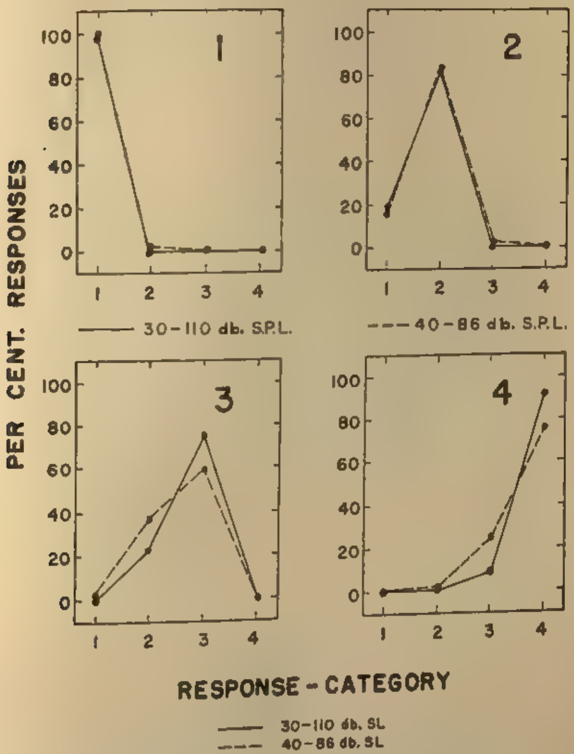


Fig. 3. The Percentage of Trials on Which the *Ss* Used Each of the Four Categories of Response

(The larger number in each group represents the four stimuli used.)

be noted in Fig. 2, judgments of the 86-db. tone which served as *CS* for half of Hovland's *Ss*, were fairly evenly distributed over stimulus-categories 3 and 4. What is more discouraging is the failure to produce any great improvement in discriminability by doubling the range of intensities. Either we must limit the conclusions drawn from a test of generalization

to a comparison with imperfect discrimination, or, as the remaining alternative, we can reduce the number of stimuli employed.

Inspection of the curves found in this experiment discloses that lack of discriminability is primarily restricted to tones immediately adjacent to the one serving as a standard stimulus. Examination of typical gradients indicates, however, that effects of generalization are not restricted to those stimuli adjacent to the conditioned stimulus. Thus, although the findings of the present study are not necessarily incompatible with a Hullian interpretation of stimulus-generalization, they emphasize the need for selecting stimuli that are absolutely discriminable and indicate such contamination must be partialed out of existing generalization gradients.

Summary

Evidence is presented to indicate that the four auditory stimuli which varied in loudness and which Hovland and others have used in their studies of stimulus-generalization, although separated by 50 jnd are not absolutely discriminable. Four other tones, obtained from an equal scale of discriminability which was constructed from a range of stimuli wider than that employed by Hovland were also found not to be absolutely discriminable. Implications of these findings for stimulus-generalization studies are indicated.

EFFECTS OF CONTEXTUAL CUES ON LEARNING FROM CONNECTED DISCOURSE

SLATER E. NEWMAN, North Carolina State College, and
ELI SALTZ, Wayne State University

Although considerable information has been gathered about factors that affect learning of paired associate and serial lists, less is known about learning from connected discourse. On the one hand, many elements of simpler types of learning tasks appear to be present in connected material (*e.g.* similarity among words that are regarded as stimuli or responses) ; on the other hand, new elements appear in the connected material (*e.g.* less flexibility in the sequence of words, the role of specific parts of speech). The study reported here is the first in a series which has the object of investigating the learning of connected material and its relation to rote learning. This particular study investigates the effects of similarity among adjectives used to modify the predicate nouns of simple sentences having the form, "Trading posts are on the swift Dalton." Studies by Dulsky and by Weiss and Margolius have demonstrated that the presence of contextual cues, such as the color of the card on which the stimulus-term appears, facilitates the learning of paired-associates.[1] The effect of variations in similarity among contextual cues has not, however, been investigated. We shall treat the adjectives as contextual cues in this sense.

The effects of similarity among terms within a paragraph or among the terms of different paragraphs on learning of prose material has been investigated.[2] Cofer analyzed the errors made in memorizing prose passages and suggested that similarity might account for many of the errors. Studies employing a design of retroactive inhibition have not, however, found any effect due to similarity between original and interpolated paragraphs.

The present study was concerned with the effects of similarity of con-

* Received for publication November 18, 1958. This study was supported by a grant from the Faculty Research and Professional Development Fund of North Carolina State College. The authors' thanks are due Mr. Harry Dorsett, through whose efforts students from Meredith College served as *Ss* in this study, and Mr. James P. Jenkins, who assisted in gathering the data.

[1] S. G. Dulsky, The effect of a change of background on recall and relearning, *J. exp. Psychol.*, 18, 1935, 725-740; Walter Weiss and Garry Margolius, The effect of context stimuli on learning and retention, *ibid.*, 48, 1954, 318-322.

[2] C. N. Cofer, An analysis of errors made in the learning of prose materials, *ibid.*, 32, 1943, 399-410; J. F. Hall, Retroactive inhibition in meaningful material, *J. educ. Psychol.*, 46, 1955, 47-52; D. P. Ausubel, Lillian C. Robbins, and E. Blake, Jr., Retroactive inhibition and facilitation in the learning of school materials, *ibid.*, 48, 1957, 334-343.

textual cues within a passage of connected discourse and with determining also whether predictions about the learning of such material could be made on the basis of Gibson's theory of generalization.[3] The following predictions were tested:

Prediction 1. As similarity among adjectives serving as contextual cues is increased, the amount learned will decrease.

Prediction 2. An error will more frequently involve the appropriate response to a similar item than to a dissimilar item.

Prediction 3. As similarity among adjectives serving as contextual cues is increased, difficulty in recognizing the noun to be associated with each adjective will increase.

Prediction 4. In testing for the association of an adjective with its appropriate noun, an error will more frequently involve the appropriate response to a similar adjective than to a dissimilar adjective.

Method

Subjects. A total of 40 student volunteers enrolled in undergraduate courses in psychology at North Carolina State College or at Meredith College served as *Ss*. They were assigned randomly to a High-Similarity Group ($N = 20$) or a Low-Similarity Group ($N = 20$).

Materials. The following paragraph was prepared in both a High-Similarity version and a Low-Similarity version.

The Rivers of Madagascar

Each of the rivers of Madagascar has a single distinguishing feature. Trading posts are on the (1) Dalton. Historic sites are on the (2) Gibson. Fishing villages are on the (3) Lawson. Shipping ports are on the (4) Hebron. Industrial cities are on the (5) Triton. Forested lands are on the (6) Barton. Resort areas are on the (7) Wilton. Farming communities are on the (8) Nelson. Mining towns are on the (9) Kenyon. The island is located in the Indian Ocean off the coast of Africa.

The blank spaces were filled in with the following terms:

High-Similarity	Low-Similarity
(1) swift	(1) swift
(2) fast	(2) clear
(3) rapid	(3) ancient
(4) quiet	(4) quiet
(5) calm	(5) narrow
(6) smooth	(6) deep
(7) broad	(7) broad
(8) ample	(8) distant
(9) wide	(9) shallow

Each of these adjectives occurs at least 22 times per million words according to the Thorndike-Lorge general word count.[4] For the High-Similarity terms the triads

[3] E. J. Gibson, A systematic application of the concepts of generalization and differentiation to verbal learning, *Psychol. Rev.*, 47, 1940, 196-229.

[4] E. L. Thorndike and Irving Lorge, *The Teacher's Word Book of 30,000 Words*, 1944.

composed of Items 1, 2, and 3; 4, 5, and 6; and 7, 8, and 9 are synonyms for one another.[5] In the Low-Similarity terms, this is not the case.

Learning. Each *S* was given five 2-min. periods to read aloud one version of the paragraph. If *S* finished the paragraph before the 2-min. period was up, he started reading it over again. Each *S* read aloud at his normal rate. *E*, using a stopwatch, timed *S*'s reading rate and recorded it. After each of the first four 2-min. periods, a test covering material presented in the paragraph was administered to *S*. After the final reading period, Tests 5 and 6 were administered.

Testing. Each of the first five tests consisted of nine items, each presented on a 5×8 card. An item contained the stem of one of the nine sentences, but with the adjective and the name of the river omitted. On the right side of the card, the names of each of the nine rivers were presented in random order. Thus a typical card might be:

Shipping ports are on the: Kenyon, Hebron, Gibson, etc. During each of the tests, *S*'s task was to read aloud the stem of the item and then to select the name of the river with which it appeared in the paragraph.

The test-items were arranged in three different orders as the first, second and third tests administered to each *S*. Tests 4 and 5 repeated the orders of Tests 1 and 2. In addition, the names were listed in three different orders on the cards on each of the first three tests. Neither the order of the items nor the order of the names of the rivers on the cards coincided with the order in the sentences of the original paragraph.

Items for Test 6 consisted of an adjective on the left of a 5×8 card and the names of all nine rivers on the right. There was one item for each of the nine adjectives for each group.

The items were presented by *E* at a 5-sec. rate. The interval between the end of the learning period and the beginning of testing was approximately 15 sec., between the end of testing and the beginning of the next learning period, 10 sec., and between the end of Test 5 and the beginning of Test 6, 15 sec. *E* recorded all responses made by *S*.

RESULTS

Prediction 1. Effect of similarity of context. As predicted on the basis of the theory of generalization, the *S*s who learned material containing similar adjectives learned more slowly than the *S*s who learned material containing dissimilar adjectives. Using the data from Tests 1 to 5, the difference between the two groups in number of correct responses was significant at beyond the 5% level.[6] The greatest difference, as can be seen in Table I was obtained at Test 3 ($p < 0.01$), with sizeable differences at Tests 2 and 5 ($0.10 > p > 0.05$). Number of correct responses using each of the nine response terms was also, in each case, higher for the Low-Similarity condition.

Prediction 2. Effect of similarity on pattern of errors. The data analyzed

[5] P. M. Roget, *Thesaurus of English Words and Phrases*, 1941.

[6] Since Bartlett's test indicated extreme heterogeneity of variance ($p < 0.001$), the Mann-Whitney *U*-test was employed in analyzing the data.

were again from Tests 1–5. For the High-Similarity *S*s, a count was made of the number of times responses for Dalton, Gibson, and Lawson, for Hebron, Triton and Barton, or for Wilton, Nelson and Kenyon were substituted for one another. These are referred to as intra-triad errors. Similar values were computed for *S*s of the Low-Similarity group. Only on Trial 1 were there significantly more ($p < 0.05$) intra-triad errors for High-Similarity than for Low-Similarity *S*s.[7] Similar results were obtained using the ratio of intra-triad errors to total errors. The number of inter-triad errors (*i.e.* errors made by substituting a response from another triad) was significantly higher for High-Similarity *S*s than for Low-Similarity *S*s ($p < 0.01$) only on Trial 3. Observed differences between the two groups in number of omissions were found not to be statistically significant. Thus the pattern of errors is not strikingly in accord with prediction on the basis of a theory of generalization.

TABLE I
PERFORMANCE ON EACH TEST BY FOUR CRITERIA
(Total possible on each test was 180; 9 responses by 20 *S*s)

Performance	Trial						
	1	2	3	4	5	1–5	6
Number correct:							
High-Similarity	19	26	35	59	72	211	36
Low-Similarity	18	45	78†	81	101	323*	61*
Intra-triad intrusions:							
High-Similarity	47*	31	29	32	21	160	32
Low-Similarity	27	42	29	28	19	145	42
Inter-triad intrusions:							
High-Similarity	103	106	105†	79	77	470	106*
Low-Similarity	124	89	65	69	56	403	72
Omissions:							
High-Similarity	11	17	11	10	10	59	6
Low-Similarity	11	4	8	2	4	29	5

* Difference between High- and Low-Similarity groups, significant beyond the 5% level of confidence.
† Difference between High- and Low-Similarity groups, significant beyond the 1% level of confidence.

Prediction 3. Effect of similarity of context on association between an adjective and its noun. A Mann-Whitney test was applied to the data of Test 6 (in which *S*s were presented with each adjective and required to respond with the appropriate noun). The tendency for Low-Similarity *S*s to make more correct responses than High-Similarity *S*s was significant beyond the 5% level.

[7] Mann-Whitney tests were employed for all comparisons.

Prediction 4. Effect of similarity on the pattern of errors: Test 6. The number of intra-triad errors on Test 6 was *lower,* though not significantly so, for *S*s who learned the more similar material than for *S*s who learned the less similar material ($p > 0.05$). The number of inter-triad errors was, however, significantly *greater* for the High-Similarity Group ($p < 0.05$). The proportion of such errors was also greater in the High-Similarity condition. No differences were observed between the two groups in number of omissions.

Thus Predictions 1 and 3 based on a theory of generalization are supported by the data, Prediction 2 is partially supported, and Prediction 4 is not supported.

Speed of reading. To determine if the differences reported above were associated with differences in speed of reading, an analysis of variance for reading speed (*i.e.* number of seconds to read the material the first time through on each of the five trials) was performed. There were no significant differences in reading rate between the two experimental groups, either in total scores or on a single trial. The F for variance due to trials was 9.52 ($p < 0.001$), suggesting that as the number of trials increased, the speed with which *S*s read the paragraph increased.

Discussion

Increasing the similarity among the contextual cues within a paragraph appears to result in (a) increased generalization mainly among similar terms during early learning trials, and (b) increased generalization throughout the paragraph during later learning trials. At the same time, less is learned from the paragraph.

A theory of generalization would predict that as similarity within a paragraph is increased the amount learned will decrease. This happened. On the basis of a theory of generalization the reduction in performance would result mainly from a tendency for the responses to similar terms to be substituted for one another. Thus, the number of intra-triad errors should be greater for the more similar material than for the less similar material. This expectation was to some extent supported by the data dealing with the effects of similarity among contextual cues, especially on Test 1. It would also be anticipated, however, that the tendency to generalize between items of *different* triads in the High-Similarity material would be little affected. The inter-triad data did not support this prediction. Thus, subsequent to Test 1 there were both a greater number of inter-triad errors, and a greater proportion of such errors by the High-Similarity *S*s than by the Low. That these results were probably not a function of omitting the

adjectives during testing might be inferred from the results of Test 6 in which number of correct responses by High-Similarity *S*s was lower than by Low, though the number of inter-triad errors and the proportion of these were higher. Whether these results for inter-triad errors on Tests 2–5 would have been obtained had recall been used for measuring retention, had a different technique for measuring similarity been used, or had the items within a triad been less contiguous cannot be answered from these data.[8]

The results of this study suggest that the concept of generalization may be useful in predicting learning from connected discourse; but the concept of generalization, at least as employed in Gibson's theory, does not alone explain all of the present results. We must therefore conclude that (a) Gibson's concept of generalization must be modified, or (b) that additional processes must be identified which, together with generalization will explain our results, or (c) that the concept of generalization must be rejected and other processes identified which may handle these results more parsimoniously. These conclusions are similar to some that we have reached elsewhere.[9]

---

[8] R. M. Gagne, The effect of sequence of presentation of similar items on the learning of paired associates, *J. exp. Psychol.*, 40, 1950, 61-73; S. E. Newman, Effects of contiguity and similarity on the learning of concepts, *ibid.*, 52, 1956, 349-353.

[9] S. E. Newman and Eli Saltz, Isolation effects: Stimulus and response generalization as explanatory concepts, *J. exp. Psychol.*, 55, 1958, 467-472.

THE PERCEPTION OF RADIATION BY ALBINO RATS

By LOREN C. LOGIE, JOSEPH S. PIZZUTO, Brooks Air Force Base,
JOHN E. OVERALL, and W. LYNN BROWN, University of Texas

This study is the fourth in a series dealing with the behavior of albino rats during prolonged periods of exposure to X-radiation of moderate intensity.[1] Each of the earlier studies was undertaken (1) to duplicate previous findings and (2) to examine additional hypotheses.

Previous investigations have indicated that X-radiation reduces the time that rats spend in a normally preferred surrounding and that this effect is more pronounced among female than male rats.[2] This behavior is apparently due to a generalized reaction to radiation induced by physiological alterations and not dependent upon the rats' ability to discriminate between radiated and non-radiated areas. The results of the earlier studies did indicate, however, that radiation may be perceived and responded to as a noxious stimulus. Two rats, which were confined in shuttle-boxes with two identical black compartments where one compartment was exposed to radiation and the other compartment was shielded from it, spent only 19% of the time in the exposed area. Using a pooled estimate of error derived from the complete experiment, we found that this value differed significantly from the 50% which would be expected if there were no basis for preference between the two compartments.

EXPERIMENT I

Experiment I was designed to answer several specific questions: (1) When, given a choice between two otherwise identical dark compartments, can albino rats distinguish between compartments exposed to and shielded from radiation? (2) If discrimination is possible, do male and female rats differ in responsiveness to the stimuli? (3) If sex differences are ob-

* Received for publication August 4, 1959. This work was conducted at the Radiobiological Laboratory, operated jointly by The University of Texas and the USAF School of Aviation under Contract AF 41(657)-149.

[1] F. H. Rohles, J. E. Overall, and W. L. Brown, Attempts to produce spatial avoidance as a result of exposure to X-radiation, *Brit. J. Radiol.*, 32, 1959, 244-246; J. E. Overall, W. L. Brown, and L. C. Logie, Instrumental behavior of albino rats in response to incident X-radiation, *ibid.*, 32, 1959, 411-414; Overall, Logie, and Brown, Changes in the shuttle-box behavior of albino rats in response to low-level X-radiation, *Rad. Res.* (in press).

[2] Overall, Brown, and Logie, *op. cit.*, 32, 1959, 244-246; Overall, Logie, and Brown, *op. cit.*, *Rad. Res.* (in press).

served, can they be attributed to a biological-hormonal mechanism which is altered by gonadectomy? This experiment was, therefore, an attempt to replicate the black-black shuttle-box condition of a previous experiment with a sufficiently large number of *S*s to establish reliability of the effect.[3]

Subjects. Forty albino rats of the Sprague-Dawley strain (10 normal males, 10 normal females, 10 gonadectomized males, and 10 gonadectomized females) were used as *S*s. The gonadectomies were performed about 3 weeks before the *S*s were used. All were in good health at the time of the experiment.

Apparatus. The apparatus, previously described,[4] consisted of a bank of four shuttle-boxes placed at distance of approximately 295 cm. in front of a Picker 260 KVP Deep Therapy X-ray machine. This distance, with radiation constants of 7 ma. and 250 KVP and with added filtration of 1 mm. Al and 0.5 mm. Cu, resulted in an average dose rate of 1 r. per min. inside the non-shielded compartment. A ¼-in. lead screen was placed 28 in. in front of one of the two identical black compartments of each shuttle-box.

Each of the shuttle-boxes was so balanced on a fulcrum that the weight of a rat in either compartment resulted in a vertical displacement of ¼-in. that operated low-friction switches of electronic timing devices which recorded the number of seconds *S* spent in each compartment. A 5-w. tungsten bulb was located 8 in. above the floor of each shuttle-box directly over the fulcrum. No other source of illumination was present during the experiment.

The shuttle-boxes (30 in. long, 5¾ in. wide, and 6½ in. high) were constructed of ¼-in. plywood. Each had a sliding lid constructed of expanded metal grating, and each was painted a flat black inside and out. A metal partition 1-in. high, located at the fulcrum, separated the compartments.

Procedure. On every experimental day one *S* was taken in turn from each of the four groups (one normal male, one normal female, one gonadectomized male, one gonadectomized female). These *S*s were then placed in planned random order in one of the four shuttle-boxes. Care was taken to vary their placement; half were placed in the right- and half in the left-hand compartments.

On half the days a lead shield was placed 28 in. in front of the left-hand compartments, and half in front of the right-hand compartments. Half of the *S*s of each group were run with the shield to the right and the other half with the shield to the left. The order of left-right placement was randomized within the 10 days required to run all the *S*s.

The *S*s were placed in the shuttle-boxes at 7 A.M. each day and the X-ray unit was kept in continuous operation for the next 20 hr. The number of seconds spent in the right-hand compartment was recorded at 30-min. intervals for each *S* during that period. The proportion of time spent in the right-hand compartment was calculated for each *S*, and the data was subjected to a three-way factorial analysis of variance.

Results. Table I presents a summary of the analysis of variance comparing time in the right-hand compartment for (a) shield-right and shield-

[3] Overall, Logie, and Brown, *op. cit.* (in press).
[4] Rholes, Overall, and Brown, *op. cit.*, 32, 1959, 244-246.

left, (b) male and female, and (c) normal and gonadectomized. Fig. 1 permits comparison of the mean proportion of time the *Ss* of the four groups spent in the right-hand compartment when the shield was on the right and on the left.

Over-all, no significant difference in proportion of time spent in the

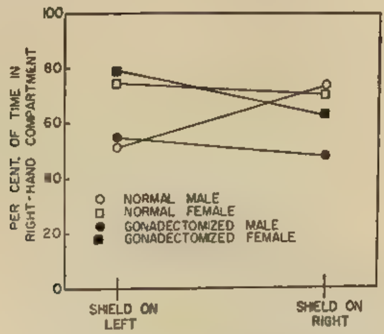


FIG. 1. PERCENTAGE OF TIME EACH GROUP OF *Ss* SPENT IN THE RIGHT-HAND (SHIELDED) COMPARTMENT

TABLE I

ANALYSIS OF VARIANCE CONSIDERING EFFECTS OF SHIELD-LOCATION, SEX, AND GONADECTOMY UPON PROPORTION OF TIME SPENT IN RIGHT-HAND COMPARTMENT

Source	*Ss*	*df.*	*ms*	*F*	*P*
Shield	14.32	1	14.32	<1.00	N.S.
Sex	2299.62	1	2299.62	10.31	<0.005
Operation	397.34	1	397.34	1.78	N.S.
Shield × Sex	746.93	1	746.93	3.35	<0.10
Shield × Operation	1097.98	1	1097.98	4.92	<0.05
Sex × Operation	238.58	1	238.58	1.07	N.S.
Shield × Sex × Operation	186.06	1	186.06	<1.00	N.S.
Error	7136.50	32	223.02		
Total	12117.35	39			

right-hand compartment resulted from differences in location of the shield. Disregarding sex, the position of the shield did, however, affect normal and gonadectomized *Ss* differently ($p > 0.05$). On the average, normal rats evidenced a tendency to move in the direction of the shield compartment, while gonadectomized rats indicated a slight tendency to move in the opposite direction.

Actually, all of the differences observed for normal rats between the two shield-placement conditions was due to the normal male group. Assuming homogeneity of variance and using the basic estimate of error var-

iance from the over-all analysis of variance, the difference in mean proportion of time spent in the right-hand compartment for the two shield-placement conditions is statistically significant for normal male animals ($p > 0.05$). Virtually no difference is observed for female rats under the two treatment-conditions. The apparent shifts in the direction of the exposed compartment by the two groups of gonadectomized rats are not statistically significant.

The over-all difference between male and female rats in response to the two shield-placement conditions fails to reach statistical significance ($p < 0.10$). Testing this effect for normal animals only and using the estimate of error from the over-all analysis of variance, we find that the difference in response of male and female rats to the two treatment conditions is statistically significant ($p > 0.05$). In this experiment, male rats indicated the greater shift in the direction of the shielded area.

A finding which makes interpretation of the above results more difficult is the over-all difference between male and female rats in preference for the right-hand compartment ($p > 0.005$). This is not considered to be a radiation effect since half of the *Ss* were exposed in the right-hand compartment and half in the left-hand compartment. In reality, every attempt was made to reduce extrinsic factors which might have produced unequal preference for the two compartments; nevertheless, each group of female rats (both normal and gonadectomized) indicated a preference for the right-hand compartment, and this preference was relatively independent of the position of the lead shield. On the other hand, three of the four groups of male rats indicated no preference between the two compartments, and the fourth group (normal males) preferred the right-hand compartment when that compartment was shielded from radiation.

Experiment II

The results of Experiment I suggested that normal, male, albino rats are capable of perceiving X-radiation. Experiment II was, therefore, undertaken to establish the reliability of these results.

Subjects. Twenty normal, male, albino rats of the Sprague-Dawley strain were used as *Ss*. They were approximately 50 days of age at the beginning of the experiment.

Apparatus. The apparatus and radiation source were the same as described in Experiment I. A bank of five shuttle-boxes was employed and each shuttle-box consisted of two identical black compartments. A lead screen was employed to shield one of the compartments from radiation.

Procedure. The *Ss* were randomly divided into two groups. Those in Group A

were placed in the shuttle-boxes with the lead screen located in front of the right-hand compartment; those in Group B, with the screen located in front of the left-hand compartment. The *Ss* were run five at a time, and the lead shield was placed on the right and left on alternate days.

The X-ray unit, as in Experiment I, was run continuously for 20 hr. and the amount of time spent in the right-hand compartment was recorded as before. The statistical analysis consisted of comparison of the proportion of time spent in the right-hand compartment by the *Ss* of the two groups.

Results. The *Ss* of Group B spent an average of 56.2 ± 30.8% of the time in the right-hand (radiated) compartment. Those of Group A, an average of 71.1 ± 18.5% of the time in the right-hand (shielded) compartment. Because of the large standard deviations, these means do not, by the usual statistical test, differ significantly. Thus, once more there is a suggestion that radiation may be perceived and responded to as a 'noxious stimulus,' but the evidence is inconclusive.

Experiment II was designed to be an exact replication of Experiment I. If the estimate of variance of population-error derived from Experiment I is assumed to represent the true population-variance, as it was in deciding upon the number of *Ss* necessary to give sufficient power to the test of significance in this experiment, the mean difference is significant.

Discussion. In each of these experiments there is a suggestion that albino rats are capable of discriminating between areas exposed and non-exposed to radiation. These results support a similar suggestion derived from an earlier experiment. The most striking observation is the weakness of the effects; with the large error-variances involved, definite conclusions concerning the existence of a perceptual mechanism capable of detecting low-intensity X-radiation would involve considerable lack of scientific caution. Because of the extreme importance of knowledge concerning such a mechanism in rats, and perhaps also in human beings, the writers believe that these findings should be tentatively presented.

It is interesting to note that the most pronounced effect observed in the present experiment was among normal male rats. This finding is to be contrasted with the finding of a more pronounced effect among normal female rats in a previous experiment.[5] In the earlier experiment, the effect was attributed to a generalized motor response to radiation induced physiologically; in the present experiment, the observed effect is more reasonably attributed to a perceptual mechanism. If the two mechanisms do exist and are jointly responsible for changes in shuttle-box behavior under the con-

---

[5] Overall, Logie, and Brown, *op. cit.* (in press).

ditions imposed in this series of investigations, then it appears likely that the general 'physiological' mechanism is more prominent among females and the 'perceptual' mechanism more prominent among males.

SUMMARY

Two experiments designed to test the hypothesis that albino rats perceive X-radiations as 'noxious stimuli' are reported. Each experiment involved two conditions which are exactly alike except for the location of a protective lead shield. For half of the *S*s, the lead shield is located in front of the right-hand compartment of the shuttle-boxes; for the other half, the lead shield is located in front of the left-hand compartment. In each experiment the results suggest that radiation, or its immediate consequence, is perceived by the rats; however, in each experiment the observed effect is very moderate.

THE EFFECT OF INSTRUCTIONS ON SIZE-JUDGMENTS UNDER REDUCTION-CONDITIONS

By Ray Over, University of Sydney

The size of an object and its distance from *S*'s eyes determine the size of the image formed on the retina. Under ordinary viewing conditions size-judgments approximate the measured size of the object irrespective of distance; there is size-constancy. Under some conditions, however, the stimulus-object is judged smaller the further it is away. Gilinsky found that, under ordinary viewing conditions, *S*s given 'objective' instructions, were able to estimate the measured size of objects accurately, whereas *S*s given 'retinal' instructions made matches approaching the visual angle.[1] It has also been found that, when a situation is reduced by the removal of cues of distance, size-estimates approach visual-angle expectations when distance is varied.[2]

It is possible, in these reduction-experiments, that all cues to distance were not eliminated and that the results obtained were a function of *S*'s attitude. The experimental arrangements used suggest this possibility. Sometimes *S* had previous knowledge of the apparatus, the testing conditions, and the alternative theories. Usually the instructions were ambiguous with 'size-equality' undefined. Retinal judgments were made easier because the comparison-stimulus was to be compared were placed at a constant distance close to *S*. With a constant visual angle, a retinal judgment would always occupy the same area of the comparison screen.

If these variables operated, may it not be that visual angle is the relation *S* will give because he can give no other, and that visual-angle responses are given when *S* is asked to give them? If this were the case, "the instruction-stimulus might be thought of as the 'selector' of a given class of discriminations"[3] and would need to be examined.

The purpose of the study reported here is to determine the effect of two sets of instructions on the estimation of size under conditions in

* Received for publication June 8, 1959.

[1] A. S. Gilinsky, The effect of attitude upon the perception of size, this Journal, 68, 1955, 173-192.

[2] A. H. Holway and E. G. Boring, Determinants of apparent visual size with distance variant, this Journal, 54, 1941, 21-37; William Litchen and Susan Lurie, A new technique for the study of perceived size, this Journal, 63, 1950, 280-282; A. H. Hastorf and K. S. Way, Apparent size with and without distance cues, *J. gen. Psychol.*, 47, 1952, 181-188; E. L. Chalmers, Monocular and binocular cues in the perception of size and distance, this Journal, 65, 1952, 415-423.

[3] C. H. Graham, Behavior, perception, and the psychophysical methods, *Psychol. Rev.*, 57, 1950, 108.

which distance-cues have been eliminated as far as possible. If the two groups of *S*s, given different instructions, differ in their size-estimates, the difference could arise because the situation was not sufficiently reduced rather than that visual-angle estimates are artifacts. If both sets of judgments approach the visual-angle, it will be demonstrated that responses based on the visual-angle are forced by the situation and are not a function of attitude. A base line will have been reached where the *S*s cannot judge measure-size, even when told to do so.

METHODS

Apparatus. The experiment was conducted in an almost totally dark room 35 ft. long. *S* was seated at a desk inside a booth with a head-rest 3 ft. from the floor. A black curtain at the front of the booth could be raised or lowered by an assistant. When the curtain was lowered, *S* could see no light, when it was raised he could see only the two stimulus-objects. Stray light from walls and the floor was trapped by screens 18 in. in front of the headrest.

Two identical, standard and comparison, diamond-shaped patches of light cast on a milk glass screen, were exposed to the *S*s by two stimulus-boxes. The patches were illuminated by a 15-w., 240-v. globe from within each box, and their brightness adjusted to 0.0457 ft.-L., as measured by an S.E.I. photometer. *E* could alter the size of the diamonds, from 9 to ½-in. diagonals, by a variable diaphram. The centers of the diamonds were placed at the height of *S*'s eyes.

The stimulus-boxes were set on tracks separated by 28°. Either could be moved independently by *E*, from 1 to 35 ft. from *S*.

Subjects. Two groups, each of 10 *S*s, were used. All were students beginning an elementary course in psychology. None had been in the experimental room before the beginning of the study. All had 20/20 vision. They were allocated to the different instructional groups at random.

Procedure. A three variable factorial design was used. The dependent variable was a size equality judgment (*SE*), and the independent vairables were: size of the standard (*SS*)—2 in., 3 in. diagonals; distance of the standard (*DS*)—15 ft., 20 ft., 25 ft.; and distance of the comparison (*DC*)—15 ft., 20 ft., 25 ft.

Instructions. Both groups were given the same preliminary instructions:

You will be led into a dark room and seated at a desk. Put your head in the headrest and keep it perfectly still when giving judgments. You will make judgments and say them aloud whenever the curtain in front of you is lifted up. You will judge whether the light on your right is larger than, equal to, or smaller than, the light on your left. After you have made your judgment the curtain will fall and you will wait until it is lifted again. Let me tell you what it meant by equal. Look at this picture.[4] You see a series of stakes here and a single stake there. To which of the series is the single stake equal in size?"

The specific instructions were as follows:

'Objective.' Imagine you are looking at a real scene and not at a photograph. By equal in size, I mean which one would give the same reading if you were to walk to it and put a tape measure on it and then walk to the single stake and measure it.

---

[4] James J. Gibson, *The Perception of the Visual World*, 1950, 184.

Do you understand what is meant by equal? A man who is 6 ft. tall is the same size when just in front of you as when a few hundred feet away. Make all your judgments using these criteria.

'Retinal.' Imagine cutting out the picture of the single stake and trying to paste it over the pictures of the other stakes. If there is an exact fit we have size-equality. Here the single stake is much smaller than any of the series. This method is the same as saying that a 6-ft. man is much smaller when far away than when close up. You will use these criteria of size in making judgments.

Instructions were given the *Ss* in a room separate from the testing-room, and they were individually conducted to the latter with eyes closed. Viewing was monocular with the right eye. The *Ss* wore earpads that they could not localize noise.

TABLE I

MEAN ESTIMATES OF SIZE-EQUALITY (*SC*)

		Standard size (*SS*)					
		2 in.			3 in.		
DS	*DC*	Visual-angle expectations	Retinal instructions	Objective instructions	Visual-angle expectations	Retinal instructions	Objective instructions
	15	2.00	1.98	1.93	3.00	3.00	2.95
15	20	2.68	2.55	2.58	4.00	3.88	3.95
	25	3.33	3.23	3.28	5.00	4.88	4.98
	15	1.50	1.48	1.48	2.25	2.18	2.20
20	20	2.00	1.90	2.03	3.00	2.78	2.93
	25	2.50	2.44	2.36	3.75	3.60	3.60
	15	1.20	1.23	1.15	1.80	1.78	1.70
25	20	1.40	1.63	1.53	2.40	2.35	2.35
	25	2.00	2.00	1.95	3.00	2.98	2.98

SS signifies standard size in inches; *DS* standard distance in feet; and *DC* comparison distance in feet.

A period of dark adaptation of 3 min. preceded testing. Each *S* gave 18 sets of judgments, that is, each combination of size, standard distance, and comparison distance—the orders being selected at random.

A variation of the method of limits was used. The variable object was set at a size midway between the size of the standard and the size which would project an equal retinal image. Then it was moved, in ½-in. steps, in the direction *S* indicated. Each series continued for two steps after *S* had reversed his judgment. This method of measurement favors neither 'retinal' nor 'objective' judgments as either would require the standard to be decreased on half the trials and to be increased on the other half.

RESULTS AND DISCUSSION

The results for both groups, given in Table I, closely approximate the visual-angle expectations. A test of significance is not possible in experiments which try to find an expected result and not a difference.[5] The differences between the expected values and the obtained values are

[5] Harold Gulliksen, *Mathematical Solutions for Psychological Problems*, 1958.

not systematic, and probably can be accounted for in terms of relatively insensitive measuring units. For example, with *SS* 3, *DS* 25, and *DC* 15, the expected *SC* value was 1.80. All *S* could do, however, was to judge that *SC* 1.50 was too small, while *SC* 2.00 was too large. In compiling results, *S*'s judgment was taken as 1.75.

Inspection of our results indicates therefore, that 'objective' instructions have not enabled *S* to make a more accurate estimate of the real size of the stimulus-object than instructions to give a judgment the very opposite to real size. There are a set of base-line conditions such that visual-angle responses must be considered as relations the organism will give because he can give no other and not as relations that are artifacts of the instructions.

An analytic—as opposed to a demonstrational—psychology of the perception of size must commence with such conditions and, by varying particular sets of experimental features, set out relationships between stimulus-complexes and size-judgments.

PROBABILITY-MATCHING WITH AN UNSCHEDULED RANDOM SEQUENCE

By Elizabeth B. Morse, Hobart College, and
William N. Runquist, Pomona College

Count Buffon, an eighteenth century philosopher and naturalist, discovered that a small rod, if dropped onto a firm surface that had a series of equidistant parallel lines marked on it, would come to rest across or touching one of the lines with a probability that is a function of the length of the rod and the distance between the parallel lines.[1] If K is the distance between the neighboring lines, and t is the length of the rod, then the probability of contacting a line can be expressed by the formula $2t/K\pi$. This relationship was utilized in the present experiment on probability-matching. Each *S* created his own random sequence by dropping a rod on a pattern of parallel lines and predicting before each drop whether or not the rod would come to rest in contact with a line. The experiment was designed to determine whether the probability-matching which has been reported by various investigators would occur with a sequence of events clearly unscheduled by *E*.[2] Transfer from this situation to a scheduled one also was studied.

Subjects. The *S*s were 32 students in introductory psychology. They were randomly divided into two groups. The 16 *S*s in the experimental group served first in the rod-dropping experiment and then, with the same sequence, in a two-choice experiment with lights. Each of the 16 control *S*s was randomly paired with an experimental *S* and given the same sequence in the latter experiment. This design gave us a baseline from which to determine whether the experimental *S*s responded differently in the rod-dropping experiment and whether there was any transfer to the light-choice experiment.

Apparatus. The experiments were conducted in a small room illuminated only by electric light. In the first part of the study, *S* was given a small aluminum rod, ⅛ in. in diameter and 3¾ in. long. Six parallel lines, at separations of 7½ in., were drawn in chalk on the floor. These lines served merely as rough indicators. The narrow cracks between the floor-boards were the real indicators of contact.

* Received for publication July 6, 1959.

[1] George Gamow, *One, Two, Three . . . Infinity*, 1947, 210-213; Edward Kasner and James Newman, *Mathematics and the Imagination*, 1940, 246 f.

[2] D. A. Grant, H. W. Hake, and J. P. Hornseth, Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement, *J. exp. Psychol.*, 42, 1951, 1-5; M. E. Jarvick, Probability learning and a negative recency effect in the serial anticipation of alternative symbols, *ibid.*, 41, 1951, 291-297; W. K. Estes and J. H. Straughan, Analysis of a verbal conditioning situation in terms of statistical learning theory, *ibid.*, 47, 1954, 225-234.

In the second part of the study, a light-board, 2 ft. square, was mounted on a cart 40 in. high that could be wheeled into position. Three 6-v., 0.15-amp. lamps were secured to the face of the board 8 in. apart. The central lamp was blue in color and slightly higher on the board. Two 6-v. dry cells were used to power the lamps and were wired to an SPDT knife-switch which, when closed in one direction, first lighted the central blue lamp and 2 sec. later the left lamp, and when closed in the opposite direction, first lighted the central lamp and then 2 sec. later the right one. A Hunter timer controlled the 2 sec. delay-periods.

Procedure. The *Ss*, one at a time, were brought into the experimental room and read the following instructions for the first part of the study.

Instructions. From a standing position and holding the rod so that it points downward from a height of about 4½ ft. from the floor, simply drop the rod. Now try it. As you see, when the rod becomes stationary it will either be across a line or between the lines marked in chalk. Before each trial, you are to predict which position the rod will take, using the terms *crossing* or *not crossing.* If there is any doubt as to whether the rod is touching a line or not, or has rolled out of the area, the trial will not be counted, but the same prediction is to be used for the next trial. Are you sure you understand the instructions? The rest of the trials will have to be conducted without conversation or other interruption. Please make a choice on every trial even if it seems difficult. Make a guess on the first trial, then try to improve your guesses as you go along and make as many correct choices as possible. We have told you everything that will happen. There are no tricks or catches in this experiment. We simply want to see how well you can profit from experience.

Any questions were answered by demonstrating how the rod should be dropped or by reading again the appropriate part of the instructions. Every *S* then performed for 200 trials. During this time, *E* recorded each prediction and the way the rod actually came to rest, *i.e.* whether crossing or not crossing. It was this random sequence of actual falls that was used in the second part of the study. At no time did *S* have an opportunity of seeing *E*'s record. After the first 100 trials, *E* offered *S* a brief rest-period. The offer was accepted only occasionally.

At the end of rod-dropping, the light-board was wheeled into position. *S* sat facing it, about 2 ft. away, and *E* sat behind the board. The usual instructions for light-guessing were read.[3] Any questions were answered by reading again the appropriate part of the instructions. For the first eight experimental *Ss*, the left light was used for 'crossing'; while for the second eight experimental *Ss* the right light was used when 'crossing' was indicated. The record of light-predictions was kept by *E* on the same form as before.

After every experimental *S* was tested, a control *S* was paired with him. The control *S* was given only the light-guessing experiment, using the random sequence generated by the experimental *S* in the first part of the study.

Results. According to mathematical theory, when the length of the rod is half the length of the distance between the parallel lines, the probability of crossing a line is about 7/22, (0.318). The final average proportion resulting from 200 trials of actually dropping the rod by each of 16 experimental *Ss* (3200 trials) was 0.350. This slightly, but significantly, higher proportion makes no difference in the analysis of the data as each *S*'s pre-

[3] Estes and Straughan, *op. cit.*, 228.

dictions were compared with his own obtained sequence of events. All raw data were reduced to average proportions for 10 blocks of 20 trials each for each *S*. Thus 7 crossings or left-lights becomes a score of 0.35. Obviously, each of the 16 random series created by the 16 experimental *S*s was slightly different, with different lengths of run, different averages per 20 trials, and different grand averages for 200 trials. The latter ranged from 0.28 to 0.48, with $SD = 0.05$.

The progression of the predictions for the experimental and control *S*s is shown in Fig. 1. Each point is a group-average for that block of 20 trials.

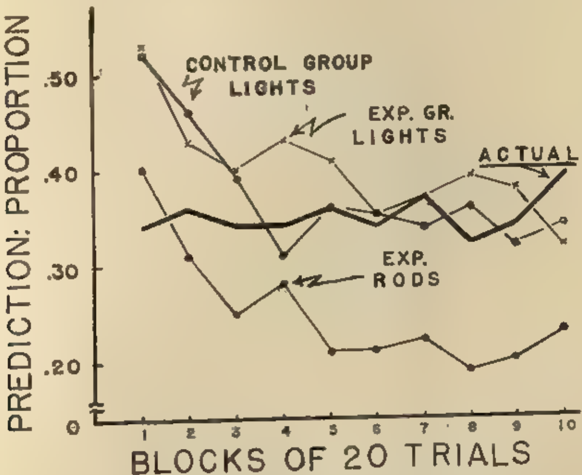


Fig. 1. Predictions Made by Experimental and Control *S*s in the Rod-Drop and Light-Choice Experiments

Each point is an average for that block of 20 trials; also shown is how the rod actually fell.

Shown also is the curve of the group-average indicating how the rod actually fell—the average of the sequences generated by the experimental *S*s. It is apparent that the experimental *S*s consistntly underestimated in their predictions regarding the fall of the rod. Both the experimental and control *S*s, however, ultimately made probability-matches, in the 'light-guessing' experiment. The similarity of the two curves of light-prediction suggests that there was no transfer, for the experimental *S*s, from the rod-dropping experiment to the light-choice experiment. Statistical analysis (*t*-tests)of the data for the last 100 trials (the first 100 trials for each *S* were assumed to be acquisition-trials) supports this conclusion.

Every *S* was questioned by *E* at the end of the study. A few of the ex-

perimental *S*s thought they could occasionally control the place the rod landed by the method used in dropping it. None of the experimental *S*s realized that the same sequence had been presented to them on the light-board. All of the *S*s (32) had tried in the light-sequences to find patterns of presentation, and, when one sub-hypothesis failed, they tried another. This result is in accord with previous findings.[4]

Discussion. Our hypothesis, that the *S*s would respond differently in a probability-situation which was more obviously unscheduled by *E*, was supported by our data. It was undoubtedly apparent to the experimental *S*s that *E* had nothing to do with the sequences in the rod-dropping experiment. They did respond in a very different manner to the rod-dropping and consistently underpredicted. Why this difference occurred is a subject for future investigation, although a possible explanation is offered by the remarks of some of the experimental *S*s. It may be that the rod just "looked short" in comparison with the distance between the parallel lines and therefore could not be expected to cross very often. In other words, it might be presumed that the experimental *S*s formed a preconceived notion of how the rod was apt to land and consistently disregarded the information given them on cumulated trials as to how it actually landed.

Perhaps an even more interesting development is that 5 of the 16 experimental *S*s maximized their predictions in the rod-dropping experiment.[5] Of the five *S*s who maximized, four announced their intention of doing so: one on Trial 15, one on Trial 32, and two on the very first trial. The decision of these last two *S*s further supports the explanation of the lower proportion in the rod-predictions on the basis of a preconceived notion that the rod was so short that it would not cross very often.

This maximizing phenomenon is a special aspect of the reaction of the *S*s to rod-dropping. It might be postulated that maximizing occurs because the *S* realizes that there is no prescheduling involved and that he himself is creating a random series, whereas, in the light-guessing experiment, all the *S*s consistently looked for patterns of presentation. It never occurred to

[4] J. J. Goodnow and T. F. Pettigrew, Effect of prior patterns of experience upon strategies and learning sets, *J. exp. Psychol.*, 49, 1955, 381-389; H. W. Hake and R. Hyman, Perception of the statistical structure of a random series of binary symbols, *ibid.*, 45, 1953, 64-74.

[5] In the original analysis of the data these scores were eliminated on the ground that they were a factor in the lower average proportions predicted for the rods. We soon discovered, however, that this made practically no difference in the over-all results because one of these *S*s had maximized in the wrong direction; the proportions being what they were, this *S*'s scores counterbalanced the scores of the other four *S*s.

the experimental *S*s that the sequence in both experiments was the same. Several investigators have also reported a tendency of their *S*s to maximize instead of match.[6]

Summary. In this study on subjective probability, a situation was employed in which the *S* generated his own random sequence of events. The results support the hypothesis that matching is related to the use of sequences which the *S* conceives to be prescheduled by *E*.

[6] S. J. Messick and C. M. Solley, Probability learning in children: some exploratory studies, *J. genet. Psychol.*, 90, 1957, 23-32; L. B. Wyckoff and J. B. Sidowski, Probability discrimination in a motor task, *J. exp. Psychol.*, 50, 1955, 225-232.

THE ROLE OF REPETITION AND SET IN PAIRED-ASSOCIATE LEARNING

By JAMES C. REED and WINIFRED D. RIACH, Wayne State University

In the learning of paired-associates it has long been assumed that repetition leads to a gradual strengthening between the items of a pair and the strength between each develops gradually with each repetition adding an increment to the bond. Rock, however, has recently postulated that associations are formed in one trial rather than by a process of gradual strengthening.[1] In a provocative experiment Rock tested his hypothesis by having a control group learn a paired association in the traditional manner. For his experimental group, however, new pairs were substituted on each succeeding trial whenever a previous pair was not learned. The number of trials to criterion and the number of errors to criterion were identical for both the experimental and control subjects. Rock interprets his findings as indicating "that repetition plays no role in the formation of associations, other than that of providing the occasion for new ones to be formed, each on a single trial."[2] The strength of Rock's position essentially rests on a failure to reject the null hypothesis. His control and experimental *S*s were, however, given different instructions. His control *S*s were instructed to *associate* the letter or letters on each card with the number, and they were told that the serial order of cards would be randomly changed from trial to trial. To eliminate a sense of surprise and lack of understanding of the task which might otherwise be expected, the experimental *S*s were told in advance that new pairs might be shown from trial to trial, although the total number would remain the same, and that their task was to *learn* all those shown at any time.[3]

Studies by Reed have shown that the influence of set is an important variable in the learning-process and the possibility remains that part of Rock's results are due to a differential set between his control and experimental *S*s.[4] The present experiment was designed to investigate the influence of set and repetition in paired-associate learning.

---

* Received for publication August 23, 1959.

[1] Irvin Rock, The role of repetition in associative learning, this JOURNAL, 70, 1957, 186-193.

[2] *Op. cit.*, 193.

[3] *Ibid.*, 187.

[4] H. B. Reed, Factors influencing the learning and retention of concepts, *J. exp. Psychol.*, 36, 1946, 71-87.

Method. Eighty students in introductory psychology were randomly assigned to four experimental groups: two levels of task, constant and variable, and two levels of instructional set. The task and presentation of that material were identical to those in Rock's study. The materials consisted of 12 pairs of letter-number combinations randomly chosen from a pool of 50. Standard procedure for simultaneous paired-associate learning was followed. On each test-trial the stimulus-letters were shown and *S* was asked to recall the number. For the constant task, the same pairs were presented on every learning-trial. For the variable task, unlearned pairs were replaced by new ones on the following learning-trial. Learning was to a criterion of one errorless trial. A metronome was used to time a 3-sec. exposure of each card and a 5-sec. interval between successive cards. Differential set was induced by instructing half of the *S*s under each task-condition to learn as rapidly as possible and the other half merely to associate the members of a pair.

Instructions. The specific instructions for the control group were:

You are going to take part in an experiment on the formation of associations. I am going to show you several cards, one at a time. On every card there will be a combination which consists of a letter or letters and a number. A trial will consist of 12 cards. When each trial is completed I will show you cards which have only the letters on them and you will tell me the number which you have associated with it. The order in which these cards are presented will be changed from trial to trial. The trials will have to be completed without any conversation or other interruptions so if you have any questions please ask them now.

The instructions for the experimental group were identical except, instead of being told to associate, they were told:

"Learn as many of these combinations as you can in each trial," and at the end of the instructions they were warned to "remember your task is to learn as many of the pairs as you can on any one trial." It should be noted that all groups were told that new pairs might appear to control for surprise.

Results. The data (error-scores) were analyzed in a 2 × 2 factorial design. The results are summarized in Tables I and II, in which the summary of the analysis of variance is presented. The main effect of task was significant ($p < 0.05$) in favor of the constant task. The main effects, the interaction of instruction and the interaction of task, were not significant.

Discussion. These findings are contrary to those obtained by Rock. The *S*s required to learn a constant set of pairs made fewer errors than those presented with new pairs for their failures at the successive learning-trials. There were, however, 4 *S*s for whom an insufficient number of pairs was available. All of these *S*s served under the variable task. For these *S*s the total number of errors made was recorded. An analysis of the errors made by the two groups on Trial 1 reveals an insignificant difference which indicates that the groups were equal in learning ability.

While the main effect of set was not significant, the difference was in the expected direction. The failure to obtain a statistically significant result may be the consequence of the lack of power of the test. It probably re-

flects the difficulty of inducing set through verbal means in a learning experiment.

A study of the data shows that Rock's results were found for two groups in this study. In Table III, together with Rock's findings, are recorded the

TABLE I

Mean Number of Errors and Trials Required to Reach Criterion

Set		Constant task	Variable task	Total for set
Learn	$\bar{X}$ errors	18.15	21.5	19.825
	s^2	56.871	96.894	77.789
	$\bar{X}$ trials	4.55	4.95	4.75
	s^2	0.892	1.523	1.217
Associate	$\bar{X}$ errors	20.55	25.7	23.125
	s^2	90.365	106.221	102.573
	$\bar{X}$ trials	5.15	5.25	5.20
	s^2	1.607	1.25	1.394
Total for task	$\bar{X}$ errors	19.35	23.6	
	s^2	73.207	103.476	
	$\bar{X}$ trials	4.85	5.10	
	s^2	1.310	1.3743	

TABLE II

Summary of Analysis of Variance for Error-Scores

Source	*df.*	Mean square	*F*	*p*
Task	1	361.25	4.125	<0.05
Instructions	1	217.8	2.486	>0.10
T×I	1	16.2	—	
w	76	87.588		

TABLE III

Mean Error-Scores for the Four Groups of this Study and the Two Groups of Rock's Study

Set	Task: constant	Task: variable
Learn	18.15	21.5 17.2 (Rock study)
Associate	20.55 17.9 (Rock study)	25.7

means for the variable task-learning group and the constant task-associate group. It can be seen that with exception of error-scores being slightly higher, the results are almost identical to Rock's including the direction of the difference. These two groups represent the replication of Rock's study, and they received different instructions.

While the results of the analysis of the error-scores support the hypothesis that repetition does aid in the formation of associations, when the

number of trials to criterion was analyzed, somewhat different findings were obtained. The effect of set or instructions was significant ($p < 0.05$), but the effect of task was not, although the difference was in favor of the *S*s under the constant task. In other words, it appears as though instructions "to learn," as opposed to "to associate," resulted in a faster rate of learning, but it did not affect the quality of their performance, *i.e.* their error-scores. Likewise the *S*s under the two tasks could be differentiated on the basis of the type of their performance but not on their over-all rate of learning. The findings and their interpretation are then in part a function of how the criterion is measured. For this experiment, the question is raised as to which variable—number of trials or number of errors—constitutes the more appropriate measurement for ascertaining the role of repetition in learning. In the authors' opinion, when the list of pairs to be learned is relatively short and the number of trials to criterion is small, then errors constitute a more descriptive and sensitive measure of learning-performance. That is, if repetition does aid in the formation of associations, it would be expected that *S*s learning a repetitive task as opposed to a variable one would make fewer errors, but the over-all rate of learning—number of trials—might not be affected due to the short list of pairs. Furthermore, since all results were in the expected direction, it suggests that increasing the length of the lists would augment the magnitude of the differences.

In summary, when the data are error-scores the results permit the interpretation that associations are formed gradually and the strength of the bond is increased until an elicited level is reached. The findings also suggest that Rock's results may in part have been the consequence of differential set in his control and experimental *S*s. If Rock's results are not a function of set, then experiments are needed to ascertain whether there are differential conditions under which associations may be formed in one trial and when repetition is a necessary factor.

SATIATION AND CO-SATIATION: A NEW METHOD

By Wirt M. Wolff, San Francisco State College

This paper describes a method for measuring satiation in the performance of a task and presents some empirical results on the reliability of the method. The method offers a purer measure of satiation than have other methods thus far described.[1] The usual co-satiation effect is curiously absent.

The traditional method for observing satiation utilizes the performance of a response over and over. *S*, himself, decides when he will quit responding. In general, he soon ceases making the original response and fails to undertake more or less analogous responses as well (co-satiation). Ideally, the tasks used in measuring satiation and co-satiation should be ones where the patterns of response are simple and differences between the tasks are not large.

As examples of tasks measuring satiation, *S* may be asked for line drawings of cats and then, for co-satiation, of rabbits; or *S* may be asked to pack spools in square boxes and then in rectangular boxes, and so on. Such tasks may be affected by the ability of the *S* to draw or to pack spools more than are tasks where *S* is asked merely to draw slanted lines at different rhythms. This was one reason Karsten preferred the latter task. The problem as to the comparability of various scores also arises, that is, the time taken is not the same as the units completed. There is also an unknown effect upon satiation that arises from *S*'s perception of his past products, and this differs among *S*s both in the number of units each has completed and in his perception of them at any particular time. Consequently, the measures may be more representative if the task does not greatly tax *S*'s ability—if the time- and unit-scores are highly related, and if *S* can see only his current product at any given time. These are the purposes of the present method.

The new method described here attempts to minimize differential ability by using a simple task, but no objective evidence is available to evaluate the success of this attempt. The present method, however, does result in highly related time- and unit-scores, and it limits *S*'s perception to his current product at any given time.

---

* Received for publication November 4, 1958. This study was conducted at Stanford University under Project NR 150-087, Contract N-onr 224(01) with the Office of Naval Research, under the direction of Dr. C. L. Winder.

[1] Anitra Karsten, Psychische Sättigung, *Psychol. Forsch.*, 10, 1928, 142-234.

Method. The apparatus used was a modified kymograph where the record-paper moved at the rate of 0.4 in. per sec. A cover was so arranged that *S* saw only that part of the paper moving under a 1-in. square opening. Thus, *S* saw only his current product. *S* was instructed to make drawings (on the paper through the opening) of a figure presented beforehand on a 3 × 5 in. card. Two figures were used: (a) two lines, 0.4 in. long, making an angle of 70° with each other, with the vertex pointing to the top of the card; and (b) an identical figure with the vertex pointing to the bottom of the card. These figures were used for measures of satiation and cosatiation, with the beginning figure being alternated from *S* to *S*.

Instructions were drawn to motivate *S*s to do well and to make clear that the *S*s were free to stop when they wished. The instructions were as follows:

Your score here will depend upon the clarity and the number of drawings you produce. Make the drawings exactly like the ones I will show you—same size and shape. You see this window here with the paper in it; now the paper will be moving, but not too fast. You are to draw figures just like the one on this card on the moving paper as long as you feel you are improving your score. Just when you feel you are not doing your score any good is a point you will have to decide. You may do this by realizing your score will depend on the number of drawings and their accuracy. It will not count for your score if some drawings are not like this. Feel free to stop when you feel it is no longer to your advantage to continue. Any questions?

All right, study the drawing and turn it over when your are ready to start. Remember, draw as many as you can and still keep all drawings like this one. Tell me when you are ready to begin. When you want to stop, take your pencil off and tell me.

Good. Now you are to continue the next part of this. Draw this figure as long as you feel you are improving your score in it. Study it, turn the card over and you may start.

The scores obtained were the time in seconds and the number of units completed under the initial conditions (satiation) and under the final conditions (cosatiation).

Results. The *S*s, college students, were 37 in number—19 men and 18 women. The correlation of time with units for the satiational task was 0.85, and the corresponding correlation for the co-satiational task was 0.92. The average number of units per second for satiation varied from 0.36 to 1.44, and from 0.36 to 1.77 for co-satiation. Incidentally, rate on the first task correlated 0.79 with rate on the second task—so these measures have at least half of their variance in common. Rate was not predetermined by the speed of the moving paper and it did differ significantly between satiation and co-satiation. The time- and unit-scores are very highly related, despite individual differences in average rate of performing the task.

A further estimate of the reliability of the task may be made by considering the correlations for time and for units completed between corresponding scores on satiation and on co-satiation. The correlation for time-score was 0.87; that for unit-scores was 0.83. (It may be recalled that for rate it was 0.79.) These results indicate acceptable estimates of the reliability of the method.

The final question arising is whether *S*s differed significantly on the

scores between satiation and co-satiation. The time-scores for satiation had a mean of 107.5 (*SD* 72.5); co-satiation had a mean of 104.7 (*SD* 77.7). The unit-scores for satiation had a mean of 77.2 (*SD* 53.9); co-satiation had a mean of 82.1 (*SD* 66.0). The correlated means of time-scores were not significantly different ($p > 0.05$) between satiation and co-satiation, nor were the means of unit-scores significantly different. The results demonstrate individual differences to a desirable extent. Although the *S*s (on the average) did take a little less time in co-satiation compared to satiation, it was unusual that *S*s completed a few more units (on the average) in co-satiation than they did in satiation. None of the differences between conditions was statistically significant.

Perhaps satiational effects do not bear the same relations to co-satiation when the *S*s see only the current product, in contrast to seeing all previous products. If co-satiation is an accepted phenomenon then the present task does not yield this effect in the usual sense. The *S*s did indeed break off the task, and by this criterion the method may be regarded as highly reliable for observing satiation alone. The experimental conditions do seem characteristic of the situation in which satiation has been exhibited, but it certainly differs in that the co-satiation effect was not demonstrated. It may be that co-satiation is a phenomenon much more dependent on situational factors that is traditionally assumed.

'GOALS' AND 'VALUES'

By DANIEL N. WIENER and DANUTA EHRLICH, VA, Fort Snelling, Minn.

The experimental study of personal values in life, particularly as they affect psychotherapy, is in a confused state. Researchers in the field are moving in different directions without standardized tools. There is no commonly accepted definition of 'values,' and one of the most troublesome obscurities in this area stems from the confusion of the terms 'value' and 'goal.'

If these two terms are synonymous, many current divisions in the literature and research become meaningless. If they are different, then the differences should be so clearly defined that the measuring instruments are not constructed, as in one major experiment in the field, "on a hunch basis" with "no particularly detailed and well-thought-out reasons" on the grounds that certain items "seemed somewhat interesting."[1] These terms have come to be used synonymously. What is considered to be an important 'goal' is also thought to be of great 'value.' To study the question systematically, the two concepts were evaluated by means of the sematic differential technique of Osgood *et al.*[2]

This technique permitted us to determine whether therapists and patients apply important descriptive words in the same strength to these terms. For example, are these concepts equally 'permanent,' equally 'unselfish,' equally 'practical'? By using this technique, we can see where 'values' and 'goals' stand, not only relative to each other and to adjectives of special significance to them, but also relative to terms studied elsewhere.

As a control of the possibility that a direct comparison of the two terms would force the rater to make artificial distinctions only for purposes of this study, two 'control' concepts were added to the two 'experimental' for rating on the semantic scales. Both of these were taken from the Minnesota Atlas.[3] They were 'tornado' and 'pigment.' The first was selected because of the extremely negative mean ratings by which it was characterized; the second, because of its relatively neutral tone, as indicated by its semantic differential profile. The four concepts were presented in a systematically varied order, and were rated on a seven-point scale for each of the descriptive words.

Two sets of descriptive words were used. One (Fig. 1) was taken from the Min-

---

* Received for publication October 8, 1959.

[1] Letter from the project director.

[2] C. E. Osgood, G. J. Suci, and P. H. Tannenbaum, *The Measurement of Meaning*, 1957, 77.

[3] J. J. Jenkins, W. A. Russell, and G. J. Suci, An atlas of semantic profiles for 360 words, this JOURNAL, 71, 1958, 688-699.

nesota study, and consists of 20 scaled words which were selected following factorial analysis as representative of semantic space generally.[4] The other (Fig. 2) consists of 20 adjectival scales selected by clinical judgment because of their apparent relevance to 'values' and 'goals.' Half of the judges rated each concept first

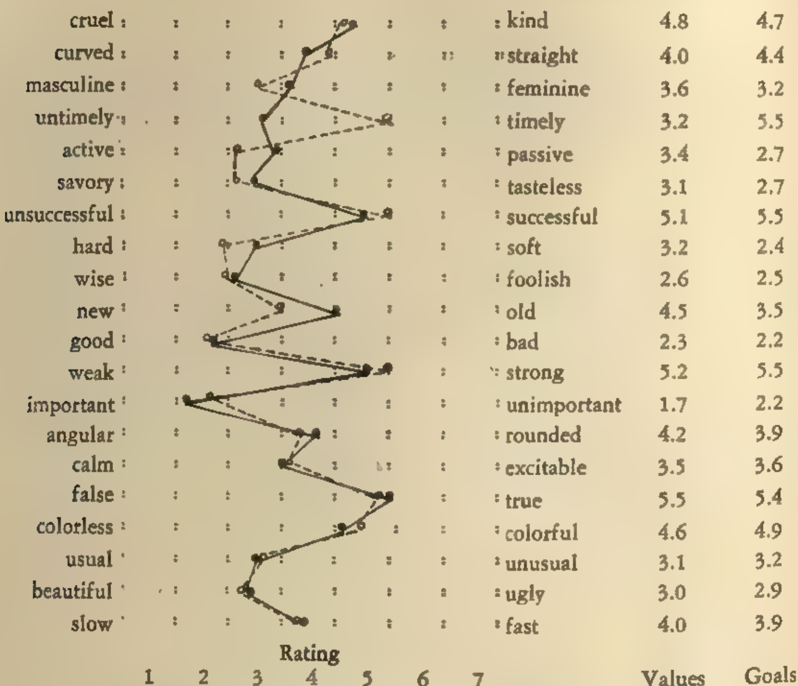


FIG. 1. RATINGS OF 'GOALS' AND 'VALUES' ON REPRESENTATIVE SEMANTIC SCALES BY 36 *Ss*:20 THERAPISTS AND 16 PATIENTS

on the Minnesota study scales, and second on the clinically-selected scales. For the other half, the order of the scales was reversed.

Twenty staff members of the clinic and 16 patients (unselected except for being in group therapy) participated in the study. The instructions for administration of the rating scales were adopted from the Minnesota study.

The mean profiles for 'Values in Life' and 'Goals in Life,' based on the ratings of the entire sample, are shown in Fig. 1 and Fig. 2. With very few exceptions, the concepts received almost identical mean ratings on both sets of scales. A separate consideration of the ratings from the staff and from the patients did not improve the discrimination between the concepts.

[4] Jenkins, Russell, and Suci, *op. cit.*, 690.

None of the scale-differences was large enough to be statistically significant.

The figures graphically demonstrate the similarity with which the two terms are described by both broadly selected adjectives and by clinically-oriented adjectives. Both 'goals' and 'values' are rated most strongly as:

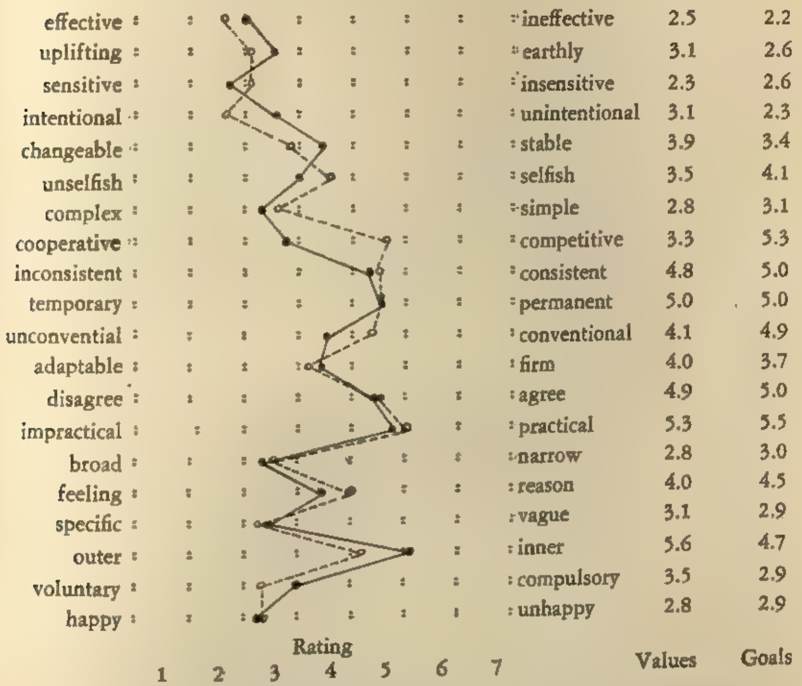


FIG. 2. RATINGS OF 'GOALS' AND 'VALUES' ON CLINICALLY-ADJUDGED SCALES BY 36 Ss:20 THERAPISTS AND 16 PATIENTS

timely, successful, strong, true, important, practical, and inner—by both therapists and patients.

In view of the similarity in the profiles for the two concepts produced by our two subsamples, and of the total, we believe that the two concepts, 'values' and 'goals,' have similar connative meaning and are used synonymously in practice. For the effective future study of 'values,' further attention to their definition and differentiation from 'goals' would appear to be useful.

AN APPARATUS FOR CONDITIONING PLANARIA

By JAMES V. MCCONNELL, PAUL R. CORNWELL, and MARGARET CLAY, University of Michigan

This paper presents a modification of an apparatus used in conditioning planaria that is more efficient and more economical than the one used in our earlier studies.[1] The present apparatus (Fig. 1) overcomes, as we believe, many of the faults in the earlier design. Instead of using a plastic tube cut longitudinally in half, we now mill a U- or V-shaped trough 10-in. long, ½-in. wide, and ¼-in. deep in a block of white, opaque plastic. Brass buttons, (B), mounted at each end of the trough, serve as electrodes which provide a current-flow of relatively uniform density throughout the trough. These buttons are connected with binding posts (see Fig. 1) and thence lead to the inductorium.

The plastic block is securely fastened to a heavy wooden base by means of an adjustable metal bolt which locks the block in place (see Fig. 2). Loosening the bolt frees the trough for cleaning or replacement. When the wooden base is placed on a laboratory bench, incidental vibration is cut to a minimum.

We find it a saving in time to condition two planaria simultaneously. Since our animals are usually given an inter-trial interval of from 30 to 60 sec., one *S* may be resting between trials while the second is being used. For this reason, we now mount two plastic blocks side by side as shown in Fig. 3. A wooden divider 12-in. high separates the two troughs. Spillage of light from one apparatus to the next is quite minimal, and the two troughs are close enough together that *E* can quite easily handle both *Ss* at once.

With conditions such as these, the number of inductoria and power-sources can be held to a minimum. For a power supply, we use an AC-DC variable low voltage power supply unit (No. 2531, produced by the GM

---

* The apparatus described here was developed with research funds granted to the senior author by the National Institute of Mental Health and by the Atomic Energy Commission.

[1] Robert Thompson and J. V. McConnell, Classical conditioning in the planarian, *Dugesia dorotocephala*, *J. comp. physiol. Psychol.*, 48, 1955, 65-68; J. V. McConnell, A. L. Jacobson and D. P. Kimble, The effects of regeneration upon retention of a conditioned response in the planarian, *ibid.*, 52, 1959, 1-5.

laboratories in Chicago), which we set to deliver 6-v. DC to a single inductorium (No. 308, Harvard Apparatus Co.). The inductorium is connected to both troughs by means of a switching arangement that allows current to flow through only one trough at any given time (see Fig. 4). In this fashion, one power supply unit and one inductorium can supply current to two (or even more) troughs.

To simplify the experimental procedure, the three switches needed to

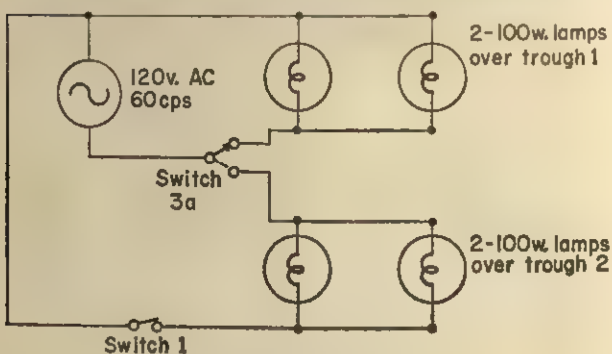
FIG. 1. THE PLASTIC TROUGH

FIG. 2. VIEW OF THE LEFT HALF OF THE APPARATUS

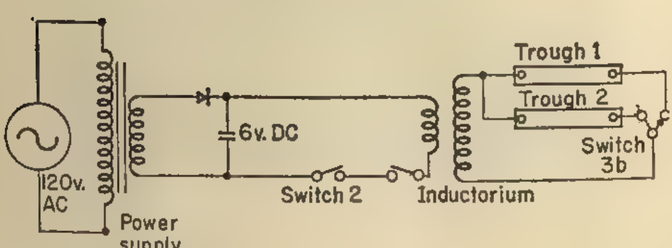
control light and shock for both troughs are mounted on a single, portable control panel (CP in Fig. 3). In the circuit diagrams shown in Fig. 4, Switch 3 is a double-pole, double-throw master switch whose position determines which of the two troughs will be operant. Switches 1 and 2 are microswitches which control the presentation of the light and the shock. When Switch 3 is positioned to the left, pressing Switch 1 turns on the lights above the left trough, while pressing Switch 2 causes a current-flow through the left trough. When Switch 3 is turned to the right, Switches 1

and 2 activate light and shock for the right trough. The portable control panel is held in *E*'s lap while he is running the *Ss*; thus, the incidental

FIG. 3. FRONTAL VIEW OF THE APPARATUS



LIGHT-CIRCUIT



SHOCK-CIRCUIT

FIG. 4. DIAGRAMS OF THE LIGHT- AND THE SHOCK-CIRCUITS

vibration from turning the switches on and off is not transmitted to the experimental apparatus.

Our studies of planaria have yielded several items of information which may be of interest to anyone experimenting with these animals.

(1) No other means of providing current to the animals is as effective as an inductorium. When other mechanisms are used (such as the well-known Harvard Stimulator), the intensity of the current needed to provoke a clear *UCR* is usually so great that the stimulation is often latently fatal. In like fashion, AC is much more deleterious than is DC. Worms shocked with sufficient AC to evoke a contraction often do not survive.

(2) Planaria can be conditioned to a decrease in illumination as well as to an increase. Here the lights are left on during the entire inter-trial interval, the *CS* being the decrease in illumination caused by switching off the experimental lights. We have found that the regular 100-w. bulbs, when left on during the inter-trial interval, heat considerably the small amount of water in the trough. This rise in temperature causes the worm to stop swimming. If a small fluorescent lamp is used instead, the problem of temperature change is avoided.

(3) The *shape* of the trough is an important variable. We have used both U- and V-shaped troughs with considerable success. Rectangular troughs make is very difficult to observe the animals when they crawl on the sides of the trough, hence we have avoided their use. Rectangular troughs are particularly to be avoided in maze-studies. When placed in U-shaped troughs, worms tend to crawl on the bottom; when the trough is rectangular, they tend to crawl on the sides. A worm crawling on the side of a rectangular trough will not turn left or right when it reaches a choice in the maze; rather, it will have to turn up or down. On the next trial, it may be crawling on the opposite side of the trough, thus its prior up-down relationships will be reversed.

(4) Planaria are best kept in water from their native habitat. If the animals are not obtained locally by *E,* almost any pond or river water is preferable to tap water. Distilled water should be avoided, since it usually causes the worms to disintegrate. Since most naturally-occurring water contains considerable mineral salts, we have never found it necessary to add electrolytes to the trough water to get a good current-flow through the troughs. Planaria are hypersensitive to chemicals dissolved in their water, and the addition of most electrolytic agents to their water might have a deleterious effect on them.

(5) Planaria are omniverous, being both scavengers and predators. Yet, while they will eat almost anything, they will flourish on only a few substances. We have had our best success with fresh-hatched shrimp eggs, which can be purchased from almost any pet shop. Particularly to be avoided as a food is egg yolk, which the worms eat avidly but which apparently does not contain the materials necessary for their growth and re-

generation. Planaria can also be maintained on daphne, mosquito larvae, small fish such as guppies (if one wounds the fish before feeding them to the worms), or fresh-water shrimp. A small planarian will attack, kill and devour a fish many times its size and weight. Many planaria will also cannabalize smaller members of their own or other species. One further word of caution: If the animals are kept in individual small bowls during training, the water in the bowls should be changed a few hours after each feeding. Otherwise, the surplus food will contaminate the water and kill the animals.

(6) Planaria, like many other species of animals, are quite sensitive to seasonal changes, even when they have been kept in the laboratory for a period of weeks or months. We have found that, at least in colder climates, planaria do not condition as readily in the winter months as they do in the spring, summer, and early fall. The seasonable differences in conditionability is often marked, hence *E* should schedule his experiments accordingly.

A CLASSICAL CONDITIONING TECHNIQUE FOR SMALL AQUATIC ANIMALS

By J. L. Horner, Nicholas Longo, and M. E. Bitterman,
Bryn Mawr College

The apparatus described here was designed for the study of classical conditioning in the fish, but it may be adapted to the study of classical conditioning in a variety of small aquatic forms, such as the crab or the newt. Inspired by the work of Bull, it measures the generalized response of an unharnessed *S* to shock and to stimuli associated with shock.[1] It consists (1) of a conditioning situation proper, which is analogous in many respects to the tambour-mounted activity-cage used in early studies of classical conditioning in the pigeon and the rat,[2] and (2) of an amplifying-integrating circuit, which provides a quantitative index of the magnitude of response, and which may be used in conjunction with any of a number of different activity-detectors.

The conditioning situation is sketched in Fig. 1. The animal's chamber (4.5 in. long, 2 in. wide, and 4 in. high) is erected on a plastic base resting on the slate bottom of a small aquarium, which itself rests on a bed of foam rubber. The narrower walls of the chamber and its top (a removable cover) are made of clear plastic. One of the longer walls (the nearer one of the sketch) is made of ¼-in. wire mesh and serves as one of the shocking electrodes. The other long wall is made of ¼-in. clear plastic slats set ¼in. apart, and beyond it is a second electrode (of stainless steel) which is fixed to the plastic base. (Fortunately—since there is simply no reasonably convenient alternative—the metal electrodes seem to have no adverse effects on the health of the fish.)

Just outside the animal's chamber, about ¼ in. away from the mesh wall and parallel to it, is a paddle of thin metal (3.75 in. wide and 1.5 in. high) which is painted

---

* This work was supported in part by Contract Nonr-2829(00) with the Office of Naval Research and in part by Grant M-2857 from the National Institute of Mental Health.

[1] H. O. Bull, Studies on conditioned responses in fishes, *J. Marine Biol. Assoc. U. K.*, 15, 1928, 485-533. Bull's attempt to develop a method which could be used with a free-swimming *S* followed his failure to obtain satisfactory results with the method of J. P. Froloff (Bedingte Reflexe bei Fischen, *Arch. ges. Physiol.*, 208, 1925, 261-271) which required the attachment of a thread to the dorsal fin. For an instance of contemporary work with harnessed fish, see L. S. Otis, J. A. Cerf, and G. J. Thomas, Conditioned inhibition of respiration and heart rate in the goldfish, *Science*, 126, 1957, 263-264.

[2] G. H. S. Razran, Conditioned responses in animals other than dogs, *Psychol. Bull.*, 30, 1933, 261-324; L. H. Warner, An experimental search for the 'conditioned response,' *J. genet. Psychol.*, 41, 1932, 91-114.

black. The paddle is fixed to a metal rod about 6 in. long which is set into the needle-holder of a crystal phonograph cartridge (Astatic L-16). The cartridge, in its shielded case, is mounted on the frame of the aquarium, which is painted black and filled with water to a depth of 3 in. The aquarium has a black plastic cover on the underside of which a small 15-w., 115-v. lamp is mounted; this lamp is the only source of illumination when the aquarium is covered, and either its onset or its offset may be used as the *CS*. The activity of the fish in the chamber displaces the paddle, and the output of the phonograph cartridge is amplified and integrated to provide a measure of that activity.

The amplifier-integrator, diagrammed in Fig. 2, is fairly straightforward, but it has

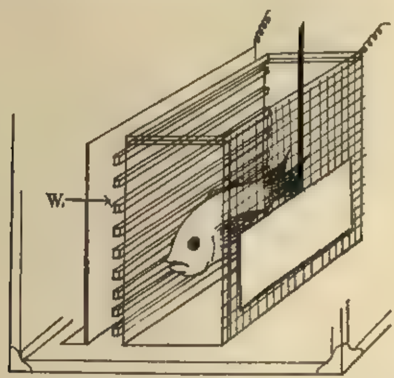


FIG. 1. SKETCH OF THE CONDITIONING CHAMBER
(*W* indicates the water-level.)

several features worth noting. Tubes V_1, V_2, V_3, and their associated circuitry constitute a voltage-amplifier with a passband of about 2–30 cycles and a maximal gain of 30,000. The choice of passband was determined by the need to detect low signal-frequencies and to suppress 60-cycle pick-up. High-gain low-frequency amplifiers tend to be unstable, and it was necessary to use voltage-integration to prevent self-oscillating. A wide-band audioamplifier is planned for more general applications.

The amplified signal is rectified by V_5, V_6, and then integrated by V_7, a high-gain, 180° phase-shift amplifier with a capacitor, C_1, connected in the feedback to make it perform as an integrator. If drift is to be kept small, C_1 must be of very high quality; in the present apparatus it is of the Stabelex-D line of the Industrial Condensor Corporation. Furthermore, all components connected at the input of the integrator must have extremely good insulation-resistance. It was necessary here to use a glass-insulated, hermetically sealed relay, RL_2, to reset the integrator. The gating relay, RL_1, and the reset-switch, S_1, both were rebuilt with polystyrene insulation for the same reason.

The output of the integrator goes to a cathode follower, V_8. This tube fires RL_2, the relay in its cathode, when the input is raised to a certain critical voltage. RL_2 resets the integrator by discharging C_1. The relay also actuates a mechanical counter. A vacuum-tube voltmeter circuit, V_4, measures the integrator-voltage and can be read

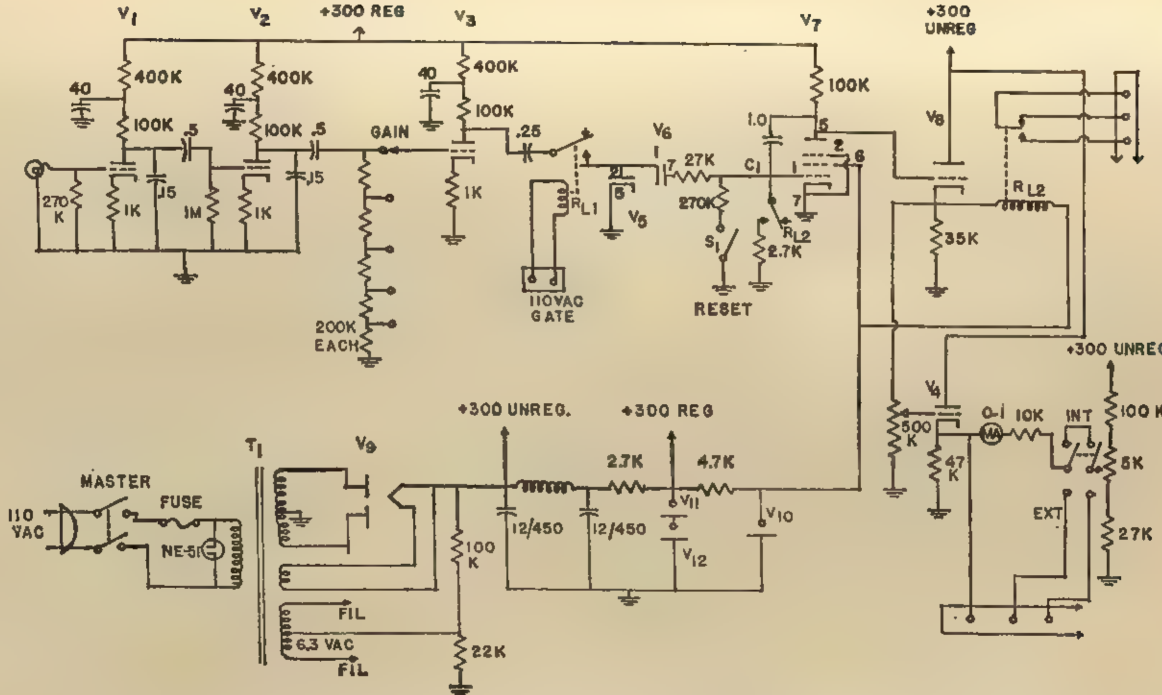


FIG. 2. DIAGRAM OF THE AMPLIFIER-INTEGRATOR
(V_1, V_2, V_3, V_4, 12AX7; V_5, V_6, 6AL5; V_7, 6AU6; V_8, 12AU7; V_9, 5Y3; V_{10}, VR105; V_{11}, V_{12}, VR150; T_1, 250-0-250 at 90 ma., 6.2 at 4 A., 5.0 at 2 A; D_1, 115N010, Time Delay, Amperite; RL_1, 100 v. AC, DPST; RL_2, hermetically sealed, 6 v. DC; M, 0-1 ma. All resistances in ohms and all condensors in microfarads unless otherwise noted.)

from the front panel. After each trial, the integrator is reset manually by the depression of a switch, S_1, on the front panel.

This device has been in operation about eight hours per day for more than four months, and in that time it has proved to be extremely stable. It has been used in conjunction with a set of six conditioning chambers. Trials are programmed by *E*, who throws a gang-switch to connect a given conditioning chamber to the amplifier-integrator and to a timing circuit, and then depresses a switch in the front panel to initiate the trial. The timing circuit controls the presentation of stimuli and the input to the amplifier-integrator (which reads only from the onset of the *CS* until just before the onset of the *US*). For purposes of partial reinforcement or extinction, *E* may prevent shock on any given trial by throwing a switch on the front panel, and there is another switch by means of which *E* may prevent the *CS* when he is interested in sampling general activity.

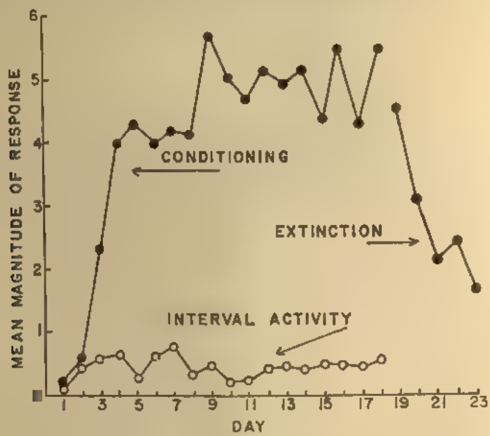


FIG. 3. THE COURSE OF CONDITIONING AND EXTINCTION IN A GROUP OF GOLDFISH

Fig. 3 shows the course of conditioning and extinction in a group of 11 goldfish trained with light-onset as the *CS*. The *CS-US* interval was 5 sec., and the termination of the *CS* coincided with the termination of the *US*, a shock of 0.5 sec. duration. The shocking voltage was 14.5 AC and the field-strength 1.45 v./cm. The current-flow was 15 ma. (rms) and the current-density 0.13 ma./sq.cm. There were 10 trials each day with an average intertrial-interval of 3 min. Fig. 3 also shows the mean intertrial activity measured during 5-sec. periods of no stimulation. The response to the *CS* of sensitization-controls given extensive experience with unpaired *CS* and *US* (10 of each per day in random order with a mean separation of 1.5 min.) did not significantly exceed their response during inter-stimulus-intervals.

APPARATUS FOR THE STUDY OF DISCRIMINATION AND CONCEPT-FORMATION

By SONIA F. OSLER, Goucher College, and MARSHALL G. POWELL, Foringer Company, Inc.

The present report describes an electronically controlled apparatus for automatically presenting two pictures simultaneously and rewarding correct choices.[1] The apparatus, which is suitable for experiments in discrimination and concept-formation, was constructed to overcome several difficulties encountered in earlier experiments with children, in which manual presentation of stimulus-objects and rewards was used.

Since the experiments required a large number of trials, the time taken with the manual procedure was unduly long. As the work progressed, additional reasons emerged for improving the procedure. The behavior and comments of some children indicated that they were competing with *E* on the assumption that he did not want them to win. They sometimes abandoned winning strategies because they expected *E* to change the rule of reward when they had succeeded several trials in a row (10 consecutive correct choices were the criterion for success). It thus became essential to remove any connection between *E* and the location of the reward. Moreover, with the manual baiting of the positive stimulus-object there was always the possibility of producing cues as to the location of the reward, despite all precautions taken to avoid them.

The apparatus described below was designed to overcome these difficulties. The pictures are projected in pairs from a film-strip and *S* presses a lever under the one of his choice. Correct choices are automatically reinforced with a marble which falls into an open cup under the levers. Following each lever-press, the shutter is closed, and, after a specified interval, the next pair of pictures is presented. Responses are recorded automatically; *E* may sit near the child or may leave the room. In either case, the competitive element between *E* and *S* seems to be eliminated.

The apparatus is designed in four units for ease in transportation to schools and other institutions where children may be tested. One unit consists of a standard film-strip projector with a single-frame attachment and a special shutter which blocks out the screen on signal. The second unit con-

[1] This investigation was supported by Grant MY-2583 from the National Institute of Mental Health to Dr. Leon Eisenberg of Johns Hopkins University and the senior author.

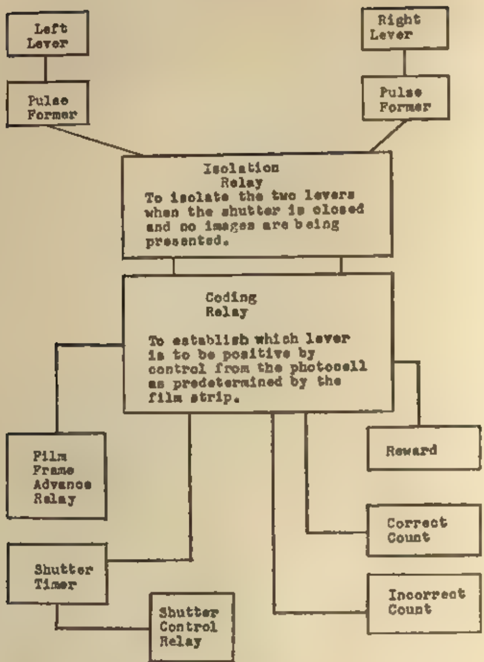


FIG. 1. SCHEMATIC DIAGRAM OF THE APPARATUS

FIG. 2. A CHILD WORKING FOR MARBLES

tains a screen which receives the projected images from the film-strip, two telegraph-keys, a container for the supply of marbles used as reinforcements, and a cup which receives each reinforcement following a correct response. The other two units house the power-supply and control equipment, such as timer and counter. Fig. 1 shows a schematic diagram of the apparatus.

The programming is automatically controlled by the presence or absence of a white spot associated with each frame on the film-strip. The length of the strip is limited only by the size of the spools. In the present studies, 150 frames with standard 35-mm. film are used. Fig. 2 shows a child working with the apparatus.

This equipment has been used with children, from 6 to 14 yr. of age, in studies of concept-formation. The mean time required for 150 trials has been reduced from 65 min. when presentation is manual to 15 min. with this automatic equipment. The children obviously enjoy the experimental sessions and even express disappointment when they are ended.

AN APPARATUS FOR THE MEASUREMENT OF SPECIFIC RESPONSE AND GENERAL ACTIVITY OF THE HUMAN NEONATE

By Lewis P. Lipsitt and Clement A. DeLucia, Brown University

The device described here was designed to measure both leg-withdrawal and generalized activity in the human neonate. It is useful for determining sensory thresholds and in the study of conditioning.

The entire apparatus is mounted in a cabinet, on the floor of which an infant's mattress is placed (Fig. 1). At the foot of the bed a completely

Fig. 1. A View of the Apparatus

enclosed section houses circuitry for photocells and light-sources. Terminal-posts for signal-output and switches for the light-sources are mounted conveniently on the outer cabinet wall.

The floor of the bed is attached to a spring-balance at one end (Fig. 2*a*), and rests on an off-center fulcrum on which it pivots. This stabilimeter-unit and the leg-withdrawal unit have independent lamp-sources (Fig. 2*c*)

* The design and construction of this equipment were made possible by Grant No. B-2356 from the National Institute of Neurological Diseases and Blindness to Brown University's Institute for Research in the Health Sciences. The writers are indebted to Dr. Glidden L. Brooks, director of the project under which this neonatal research is done, for his advice and encouragement, and to Dr. David H. Crowell, of the University of Hawaii, for a number of valuable suggestions. See also D. H. Crowell, John Peterson, and M. A. Safely, An apparatus for infant conditioning research, *Child Devel.*, 31, 1960, 47-52.

which remain stationary as their respective photocells move, thus changing the intensity of illumination falling upon them. The response-pickups de-

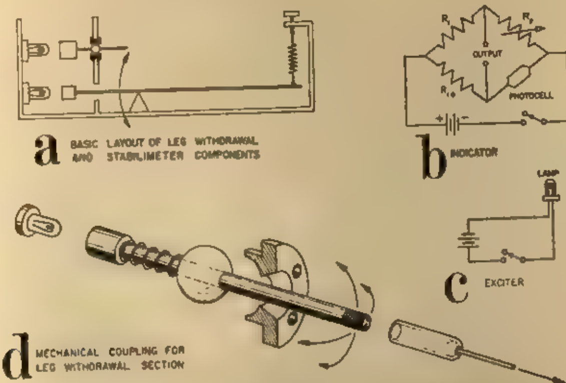


FIG. 2. COMPONENTS OF THE APPARATUS

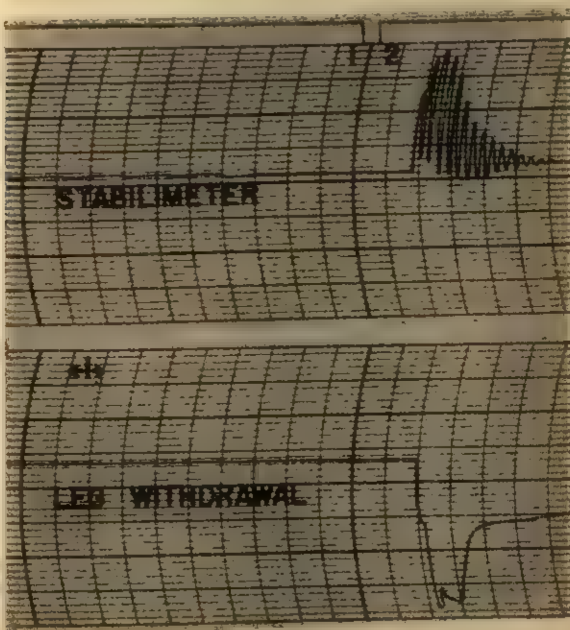


FIG. 3. SAMPLE RECORD

liver voltages which vary with the magnitude of limb- and body-movement. The basic component of each is a photocell of the resistance-varying type (Clarex Photocell Type CL 2), which is used in a bridge-circuit (Fig. 2*b*).

The leg-unit, represented in detail in Fig. 2*d*, is attached to its photocell by a spring-loaded shaft, which is attached to the wall of the cabinet by a ball-and-nylon socket. The shaft is threaded at the end opposite to that holding the photocell. A baby-shoe is attached permanently to a universal joint which may be attached to the threaded end. Thus, movement of the leg in any direction is recorded. The universal joint is attached after the baby's foot has been placed in the shoe, and is detached at the end of the experimental session.

From the terminal-posts we connect a two-chnanel Brush oscillograph (Model Mark II) which provides separate recordings of leg- and stabilimeter-movements, along with pen marked indications of onset and offset of stimuli. The type of record obtained is illustrated in Fig. 3.

Among the limitations of the present unit is the fact that the displacement of the recording pens is not linearly related to the magnitude of *S*'s movements. The device will, however, sensitively record the presence or absence of movement within a given time-interval, which is all that is needed for many kinds of threshold-determinations and for the study of conditioning.

The principles involved in the construction of this device may be adapted for the purpose of recording responses from other portions of the body.

AN IMPROVED CIRCUIT FOR THE DRINKOMETER

By Leon S. Otis and Robert A. Boenning, The Johns Hopkins University

In 1951 Hill and Stellar described an electronic drinkometer for continuously measuring the fluid intake of the rat.[1] The principle was a remarkably simple one. Each time the rat came into contact with the fluid it was to drink, it completed an electronic circuit, grounding the grid of a 6SN7 vacuum tube. Each change in grid bias caused a DC relay in series with the plate to close which, in turn, closed another circuit, activating the pens of a kymograph. Using such a device, the authors showed that the time course and pattern of successive tongue laps could be easily recorded for study.

The Hill-Stellar drinkometer is probably unsuited, however, for measuring absolute or differential gustatory thresholds or for determining fluid-preferences, because there may be interaction between the electrical circuit and the mechanism of stimulation for taste.[2] This follows because the animal completes a DC circuit in actuating the plate relay and ionization of the fluid occurs at the point of contact of the animal and the fluid. Current theories of taste place heavy emphasis on the role of anions and cations in the perception of gustatory qualities, particularly in the perception of salt and of sour.[3]

To overcome the problem of ionization of the fluid and its possible effects on taste sensitivities an improved drinkometer was designed. The schematic diagram is presented in Fig. 1 and a description of the circuit follows:

Our drinkometer consists essentially of a high-gain pentode amplifier and an electronic switch with associated power supply. The problem of ionization of the fluid is avoided by the use of an AC rather than DC actuating potential. The animal stands in an alternating electric field established between two metal plates that are connected to 115-v., 60~ A.C.[4] A short,

* Received for publication March 31, 1959. Report No. 166-I-221 under Contract N5-ori-166, Task Order I, between the Office of Naval Research and The Johns Hopkins University. Done in part under PHS Grant M-2003 from the National Institute of Mental Health to The Johns Hopkins University. The junior author is responsible for the design of the electronic circuit.

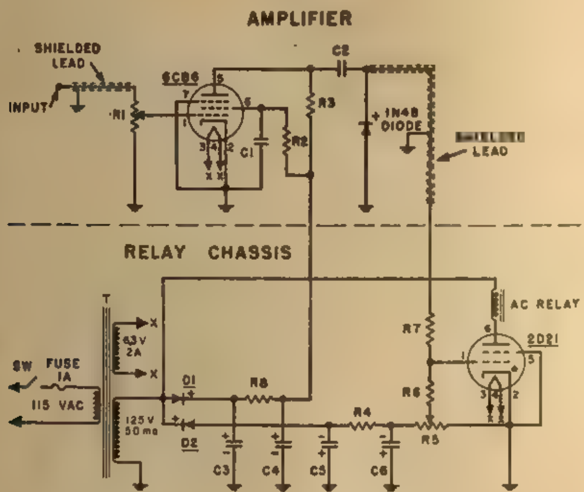
[1] J. H. Hill and Eliot Stellar. *Science*, 114, 1951, 43.

[2] Hill and Stellar did not suggest that their drinkometer would be suitable for such problems.

[3] Carl Pfaffmann, Taste and smell, in S. S. Stevens, (ed.) *Handbook of Experimental Psychology*, 1951, 1143-1171.

[4] Any good conductor which can be insulated from the animal is suitable for making plates. We used 4 × 6 in. pieces of aluminum foil cemented to the inside surface of 1/16-in. formica sheets. A pair of plates were fastened to the outside walls of each wooden drinking cage such that the rat stood between the plates while drinking. The plates were separated by about 6 in.

shielded lead connects the amplifier to a stainless steel drinking tube.[5] The input lead of the amplifier is actually an antenna to which the animal supplies a minute 60~ signal each time it licks the end of the metal drinking tube.[6] The tube is recessed slightly from the drinking cage to pre-



R1	3 MEG. POT.	C1	.05 μF, 400 V
R2	160K, 1/2 W	C2	.10 μF, 400 V
R3	33K, 1/2 W	C3, C4	40 μF, 250 V ELECTROLYTIC
R4	12K, 1 W	C5, C6	20 μF, 250 V ELECTROLYTIC
R5	10K, POT., LINEAR	D1, D2	SILICON DIODE RECTIFIERS
R6	11 MEG., 1/2 W	T	POWER TRANSFORMER, 115 V PRI., 125 V, 1/2 WAVE SEC. AT 50 ma, 6.3 V AT 2A
R7	10MEG., 1/2 W		
R8	1.2K, 1 W		

FIG. 1. WIRING DIAGRAM OF DRINKOMETER

vent the animal, physically, from applying a continuous signal to the drinking tube. The animal is required to stretch its tongue slightly in order to reach the fluid.

The signal that is applied to the drinking tube is amplified greatly and rectified by a 1N48 germanium diode. This results in a signal at the grid of the magnitude of 18 to 22 v. which consists of the positive halves of the 60~-wave. These short pulses are fed to the grid of a 2D21 thyratron tube which functions as an electronic switch. The grid of the thyratron is biased at approximately −5v. Positive B-voltage for the amplifier and

---

[5] Maximal length of shielded lead probably should not exceed 18 in.

[6] Alternately, any convenient audio-frequency at 20 v. or better may be supplied across the plates. It may be preferable to have one of the plates constitute the floor. In this case, the free lead from the oscillator is grounded to the amplifier chassis. The use of a floor-plate drops the signal requirement to a fraction of a volt.

negative voltage for biasing the thyratron grid are supplied by the two silicon diode rectifiers. A 10-kΩ-potentiometer is used to adjust the bias of the thyratron grid to a point just below cutoff. The output of the amplifier is sufficient to raise the potential of the grid above cutoff with each tongue lap and consequently, the tube conducts, causing current to flow through an AC-operated relay in the plate circuit.

The design of our circuit required the substitution of an AC-relay for the DC plate-relay of the original circuit.[7] This substitution increased the efficiency and reliability of the unit.

The circuit that we have suggested is considerably more elaborate than the original described by Hill and Stellar. Our improved design, however, has extended the utility of the drinkometer by making it suitable for use in the investigation of taste-thresholds and fluid-preferences. It has also resulted in more efficient operation, thereby inceasing the reliability of the device.

---

[7] We used a Potter and Brumfield relay KA11AY. A slight adjustment of the spring tension was found to be necessary for proper operation.

A 'SPOT REMOVER' FOR OSCILLOSCOPES

By Lloyd A. Jeffress, The University of Texas

Cathode ray oscilloscopes are often used to portray Lissajous figures which show the phase relation between two sinusoidal stimuli or the correlation between two noise signals. When the signals are turned off, a small bright spot remains in the center of the screen which often burns a hole in the phosphor. The black pit seen on the face of many a laboratory scope is mute testimony to the frequency of this occurrence.

Most oscilloscopes have an external connection to the brightening circuit. In some (Dumont, for example) this connection, labeled 'Z-axis,' appears on the front panel; in others (Hewlett-Packard) it is one of the terminals at the rear. In either case, a rather large alternating potential is needed to produce any useful change of brightness, and the signals being portrayed as a Lissajous figure are ordinarily not large enough to do this. The device to be described furnishes a voltage which is large enough and it does this only when there is an input signal. When there is no signal there is no brightening voltage and, consequently, no spot. By adjusting a threshold control on the device, one can display signals ranging from a few millivolts to several volts and still have the spot vanish before it becomes dangerously small.

Fig. 1 is a diagram of the circuit. Tube V-2 is a monostable multivibrator which remains in its stable state until the potential on its input grid (pin 7) exceeds a certain value, at which time the multivibrator oscillates. The frequency of this oscillation is determined by the values of R_{11} and C_7 and by the potential on the input grid. For the values given, this frequency is above 20 kc. If the frequency of oscillation is commensurate with the frequency to be displayed, gaps will appear in the pattern. These can be eliminated by a slight adjustment of the threshold-control and hence of the frequency of oscillation.

Tube V-1 provides two stages of voltage amplification, with a diode in the plate circuit of the second stage to furnish the additional bias needed to trigger the multivibrator. The potentiometer R_8 provides the initial bias and serves, therefore, as a threshold control for the multivibrator.

The device draws only about 1 ma. of current from the power supply. Any convenient source (250-300 v) can, therefore, be used instead of the power supply shown. If it is desired to attach the device permanently

to the oscilloscope, both the filament voltage and the plate voltage can be obtained from the power supply of the oscilloscope.

In operation, the device is connected with its input terminal to one of the signal voltages to be displayed and its output terminal to the Z-axis on the oscilloscope. To avoid by-passing the high frequency voltage, the

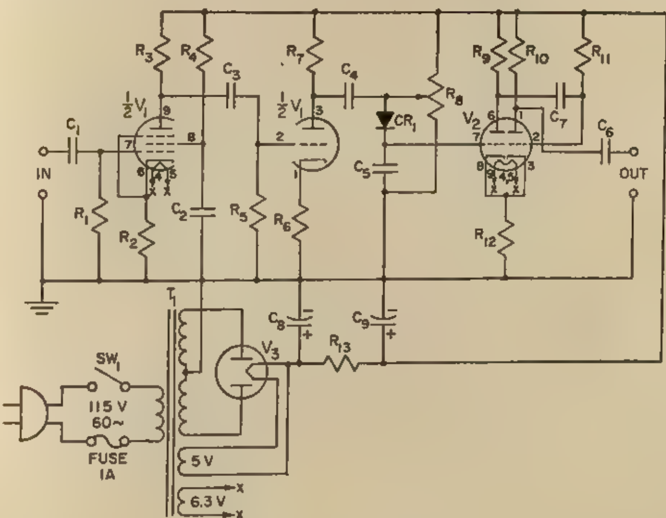


FIG. 1. CIRCUIT DIAGRAM

C_1–C_6 = 0.1 μf.
C_7 = 0.00025 μf.
C_8, C_9 = 20 μf. 450 v.
V_1 = 6AU8
V_2 = 12AX7
V_3 = 5Y3

T_1 = Power transformer 250-0-250 v., 6.3 v., 5 v.
CR_1 = 1N39B diode
R_1R_3 = 220 k.
$R_2R_{12}R_{13}$ = 15 k.
R_4 = 68 k.

R_5R_{11} = 330 k.
R_6 = 680 Ω
R_7 = 22 k.
R_9 = 33 k.
R_{10} = 100 k.
R_8 = 1 meg. pot.

connection to the Z-axis should be made by means of an unshielded wire. If this picks up too much stray voltage, a coaxial cable can be used instead. With no input signal, and with the threshold control turned down (toward ground), the intensity control on the oscilloscope is turned down until the spot is just not visible. The threshold control is now advanced until the spot is again visible and then backed off slightly until the spot again disappears. The device is now ready to operate and, when a signal is introduced, should so brighten the spot that the Lissajous figure is displayed. Small further adjustments of the threshold control will probably be needed for the particular signal voltages being employed.

NOTES AND DISCUSSIONS

HUE- AND BRIGHTNESS-DIFFERENCES, CONTOURS, AND FIGURAL AFTER-EFFECTS

Brightness-differences between figure and ground have been reported to be necessary for the appearance of clear contours.[1] Hochberg and Triebel found that hue-differences alone, between figure and ground, could neither modify nor generate figural after-effects (*FAE*),[2] but Day obtained *FAE* with figure and ground of the same brightness and different hue.[3] A series of experiments followed, with results exchanged in personal correspondence, culminating in the present set which were performed with the same apparatus, and some partial replication, in Sydney, Australia, and in Ithaca, New York.

Experiment I. Three inspection-figures were used, each being an annulus 1-mm. thick, 5-mm. in radius, spaced 12 mm. on center either side of the fixation-point. Inspection-figure *A* was red tempera-color paint, matched to Munsell 5*R* 5/6 on Gray 5/ paper of equal brightness; *B* was the same paint on 5*R* 5/6 Munsell (red) paper; *C* was black India ink on white paper. In every case, the test-figure (*T*) consisted of a pair of black annuli drawn in India ink on white paper; each annulus was 0.5 mm. thick, 4 mm. in radius, spaced 12 mm. on center each side of the fixation-point. Thus, one of the *T*-figures would appear smaller if the *FAE* occurred, since it fell on a region which was previously occupied by the *I*-figure.

There were two groups of *S*s. One ($N = 33$) was studied under Condition 1: monocular fixation with the *A*-figure at 2 ft. for 3 min., the position of the *I*-figure relative to the fixation-point being alternated randomly between *S*s; then, the *T*-figure was presented to the ipsilateral eye; this was followed by monocular fixation with the *C*-figure for 3 min., followed by the *T*-figure again. The second group ($N = 15$) was studied under Condition 2: the same as in Condition 1 except that the *B*-figure was used for the first *I*-figure instead of *A*. Illumination was dim and diffuse for both groups (0.26 ft.-L measured on the gray background paper).

[1] Susanne Liebmann, Über das Verhalten farbiger Formen bei Helligkeitsgleichheit von Figur und Grund, *Psych. Forsch.*, 9, 1927, 300-353.

[2] J. E. Hochberg and William Triebel, Figural after-effects with colored stimuli, this JOURNAL, 68, 1955, 133-135.

[3] R. H. Day, Hue differences and brightness differences as determinants of figural after-effects, *Brit. J. Psychol.*, 50, 1959, 223-230.

For the brightness-matched stimuli, *A* and *B*, the Liebmann-effect appeared, in that the contours of the annuli tended to fuzz and disappear. *FAE* appeared only for *C*, the stimulus-conditions whose figure and ground were of different brightnesses.

Experiment II. Fifteen *S*s were tested for *FAE* with inspection-figures *A*, *B*, and *C*. Here, illumination was stronger light from a window (about 63 ft.-L on the gray background of *A*). The Liebmann-effect did not appear. Contours were clearly visible, due probably to slight brightness-differences which were residual at the very edges of the tempera-color annuli incidental to their preparation. It should be noted that *FAE* was obtained with *B* which was red-on-red, *i.e.* having *neither* hue- nor brightness-differences between figure and ground. All three conditions showed *FAE*. The results of both experiments are summarized in Table I.

TABLE I
SUMMARY OF RESULTS

Condition	I-figure	Experiment I				Experiment II			
		N	*FAE*	opposed	equal	*N*	*FAE*	opposed	equal
1	A	33	13	7	13	15	13	0	2
	C	33	32	0	1	15	15	0	0
2	B	15	2	2	11	15	14	0	1
	C	5	5	0	0	—	—	—	—

Conclusion. The results of these experiments suggest that figural after-effects do not appear where figure and ground differ only in hue, if reduced brightness-differences and illumination prevent the formation of clear contours. It seems, however, that the effects will appear whenever clear contours are present, whether there are brightness-differences, hue-differences, or *neither*, between figure and ground. The conditions for contour-formation are not well explored, and the apparent dependence of *FAE* on the same conditions may provide a useful clue to the investigation of both phenomena.

J. E. HOCHBERG
R. H. DAY
D. HARDY
Cornell University
University of Sydney

AN EXTENSION OF FOLEY'S 'EXPRESSION OF CERTAINTY'

In Foley's work on the levels of subjective probability associated with the words 'sure,' 'suppose,' 'certain,' 'think,' and 'positive,' order was not counterbalanced.[1] To determine whether the results obtained by Foley were

[1] J. P. Foley, The expression of certainty, this JOURNAL, 72, 1959, 614-615.

in part a function of order, a further experiment was performed. One of five forms of a questionaire, each form with the same instructions but with the words in different serial positions, was administered to 162 men and women, students in a large section of a first-semester course in elementary psychology. In Form I, the order of the words was the same as that used by Foley and as that in which they are listed above. In Form II, the order was 2, 3, 4, 5, 1; Form III, 3, 4, 5, 1, 2; and so forth.

TABLE I
MEAN PROBABILITY-VALUES FOR EACH FORM

	Form					
	I		II	III	IV	V
	Foley (N=38)	Berry (N=30)	(N=34)	(N=34)	(N=34)	(N=30)
Positive	9.58	8.70	9.21	8.06	7.74	8.87
Certain	8.84	8.10	7.88	7.74	7.41	7.77
Sure	8.39	7.27	7.79	7.15	7.03	7.03
Think	5.97	4.20	4.00	3.94	3.94	3.97
Suppose	5.32	3.37	3.94	3.15	3.35	3.53

The results are presented in Table I, along with those of Foley. The differences in the *absolute* magnitude between Foley's data and the data presented here may be attributed to the difference between instructions, between populations from which the *S*s were drawn, and between the experimental situations. Considering such factors, the consistency of agreement with Foley as to the *ordering* of the obtained probabilities from 'positive' to 'suppose' is remarkable. The ordering is identical for all forms. The chance probabilities of obtaining such consistency in the ordering for any two words would be less than 5%; the chance probabilities for three or more words would be considerably less than 1%. Thus the present data not only substantiate Foley's ordering of subjective probabilities, but offer evidence that the ordering is stable over a variety of conditions and *S*s.

Indiana University R. N. BERRY

FILLENBAUM ON PROBABILITY-LEARNING

Fillenbaum has recently reported results obtained in a study of two-choice, non-contingent, probability-learning, which he interprets to be "somewhat at variance with the common findings."[1] After 60 trials with a 2:1 sequence of checks and pluses, his subjects (*S*s) are predicting plus

[1] Samuel Fillenbaum, Matching to objective events in probability learning: Some discrepant results, this JOURNAL, 73, 1960, 146-149.

with a relative frequency below the theoretically predicted frequency of 0.67. Fillenbaum attributes the low relative frequency of prediction to the population employed, *i.e.* Canadian Army personnel rather than college students. Had he inspected the studies on probability-learning more carefully he would have realized that learning proceeds slowly, especially as relative frequency of the to-be-predicted-event approaches 0.50, and that his series of 60 trials is far too short for the *S*s to begin to approach a stable asymptote. Although Fillenbaum's data are not reported in sufficient detail to enable statistical comparisons, inspection of his data in relation to other published data on 2:1 sequences for college students and for armed forces personnel reveals no marked discrepancy.[2] Jarvik's *S*s, it is true, show a higher initial frequency of prediction but this might well be attributable to the greater ease of writing 'check' (the preponderant event in Jarvik's study) than of writing 'plus' (the preponderant event in Fillenbaum's). There seems to be no compelling reason for treating the results as "discrepant."

New York University EDITH D. NEIMARK

THIRTY-SECOND ANNUAL MEETING OF THE MIDWESTERN PSYCHOLOGICAL ASSOCIATION

The Midwestern Psychological Association met at the Chase-Park Plaza Hotel, St. Louis, on April 28–30, 1960. Host institution was Washington University. George S. Speer, Illinois Institute of Technology, was Convention Manager, and Marion E. Bunch, Washington University, served as Chairman of the Local Arrangements Committee.

Approximately 1750 persons attended the meetings. The number of newly elected and reinstated members during the year was 317, bringing the current membership to 2455.

The program, planned by a committee under the chairmanship of William E. Kappauf, University of Illinois, consisted of 132 volunteered papers presented in 22 sessions, 5 symposia, 2 open meetings, a special workshop sponsored by the APA Board of Professional Affairs, and a meeting of Psi Chi which included volunteered papers and an invited address by Roger W. Russell.

---

[2] M. E. Jarvik, Probability learning and a negative recency effect in the serial anticipation of alternative symbols, *J. exp. Psychol.*, 41, 1951, 291-297; E. D. Neimark, Effects of type of nonreinforcement and number of alternative responses in two verbal conditioning situations, *ibid.*, 52, 1956, 209-220; E. D. Neimark and E. H. Shuford, Comparison of predictions and estimates in a probability learning situation, *ibid.*, 57, 1959, 294-298.

The papers read at the various sessions were concerned with the following topics: human conditioning, social psychology, psychopharmacology, animal behavior, subliminal perception, brain-potentials, personality, measurement and scaling, human learning and concept-formation, perception, verbal conditioning, abnormal psychology, vocational and industrial psychology, engineering psychology, reinforcement and animal-learning, attitudes and attitude-change, physiological psychology, mathematical behavior theories, child psychology, perceptual motor learning, and animal motivation.

Symposia and open meetings were concerned with: the lesser-utilized psychotherapeutic approaches and their interrelations; the psychological significance of somatic responses; the systems approach in human factors research; government programs in support of psychological research; need-strength, subjective probability of goal-attainment, and personality-assessment; chemical models for correlates of behavior; and research by psychologists in medical settings.

Ross Stagner, Wayne State University, delivered the presidential address, "Homeostasis, need reduction, and motivation." Marion E. Bunch, Washington University, was elected President for 1960-61, and John W. Cotton, Northwestern University, was elected to the Council for 1960-63. The 1961 meeting will be held May 4-6 at the Hotel Morrison, Chicago.

State University of Iowa

I. E. Farber

THIRTY-FIRST ANNUAL MEETING OF THE EASTERN PSYCHOLOGICAL ASSOCIATION

The Eastern Psychological Association met April 15 and 16, 1960, at the Statler Hilton Hotel in New York City. There were 2311 persons registered at the meetings of whom 1260 were members of the Association, 318 were new members who joined the Association at the meeting, and 733 were guests. The present active membership of the Association totals 3315.

A committee headed by Shelden Zalkind was responsible for local arrangements and the program was planned by a committee under the chairmanship of David Ehrenfreund. The program consisted of 294 papers (presented in 35 sessions), 8 symposia, 7 special meetings, 2 invited addresses, and 3 films. The 35 sessions were concerned with the following topics; visual perception (2 sessions), reinforcement, psychopharmacology (2 sessions), social behavior, comparative, brain stimulation, projective

techniques, verbal behavior, evaluation and measurement, interpersonal and self-description, animal behavior, physiological, psychopathology, learning, personality, animal brain functions, sensory processes, engineering and military, developmental and educational, human learning, decision-making, attitude structure, avoidance behavior, general clinical, operant behavior, vision, industrial and business, motivation, child, perception, thinking, brain functions, and drugs and avoidance behavior.

The symposia were titled: "Clinical implications of revisions in psychoanalytic drive theory," "Some contemporary problems in animal motivation," "Long term confinement and space flight," "Positive conceptions of mental health: implications for research and service," "Some methodological considerations in child-rearing research," "Hierarchy of personality 'layers' tapped in projective batteries," "Recent trends in special education: implications for psychological theory, practice and training," and "Differential diagnosis in schizophrenia."

In addition to special meetings entitled "Treatment of the poorly motivated patient," "Research with the Tomkins-Horn Picture Arrangement Test," and "The role of the school psychologist in the denominational school," the program also included a meeting of Psi Chi, a meeting of departmental chairmen, a meeting to discuss a proposed New England Psychological Association, and a workshop of state association officers.

Invited addresses were presented by Howard H. Kendler, "Problem solving," and by John L. Kennedy, "Growing synthetic organisms in synthetic environments." James J. Gibson presented the annual presidential address, "The concept of the stimulus in psychology."

During the business meeting it was announced that the following new officers and directors had been elected; President, S. S. Stevens, and Directors, William N. Schoenfeld and George A. Miller.

The 1961 meetings of the Association will be held at the Bellevue-Stratford in Philadelphia, April 7 and 8. The 1962 meetings will be in Atlantic City.

New York, N.Y. CARL H. RUSH

SEVENTH ANNUAL MEETING OF THE SOUTHWESTERN PSYCHOLOGICAL ASSOCIATION

The seventh annual meeting of The Southwestern Psychological Association was held at the Jack Tar Hotel in Galveston, Texas, on March 24–26, 1960. It was attended by approximately 325 members and guests.

The program contained 26 papers, 11 symposia, and the address of the

president, Ruth Hubbard, on "Some aspects of personality development in older people." Next year's meeting will be held, at the Hotel Lafayette, Little Rock, Arkansas, from April 6–8, under the presidency of Fillmore Sanford, University of Texas.

Texas Technological College BEATRIX COBB

AN ACKNOWLEDGMENT

In addition to the necrology of Mario Ponzo, the JOURNAL is indebted to Professor Leandro Canestrelli for the photograph and the signature that are reproduced in the frontispiece of this number. The photograph, a recent one, was taken in 1957 when Professor Ponzo was seventy-five years old. The signature dates from 1934.

K.M.D.

Mario Ponzo: 1882–1960

Mario Ponzo, Professor Emeritus of Psychology and Honorary President of the Italian Society of Psychology of which he had been active President in 1942-1958, died in his seventy-eighth year in Rome, after a long illness, on January 9, 1960.

Born in Milan on June 23, 1882, of a Piemontese family, he studied medicine at Turin, contemporaneously beginning his psychological training in the strict school of Friedrich Kiesow, a student of Wundt.

In 1906 Ponzo obtained his medical degree and in 1911 was appointed Docent in psychology, remaining for some twenty-five years the assistant of his master, Friedrich Kiesow.

Called to the Faculty of Medicine, University of Rome, in 1931, Ponzo succeeded Sante De Sanctis in the chair of psychology after the year in which it was held by the latter's pupil, Ferruccio Banissoni. Ponzo held this post and directed the attached Institute until 1952, when he reached retiring age for a staff-professor. He continued for the usual additional five years as a professor of psychology not on the staff. In 1958 he was given the title of Professor Emeritus.

From 1931 to 1952 he was also lecturer in psychology at the Faculties of Philosophy and Law.

He was a member of various non-Italian psychological societies, a foreign associate of the American Psychological Association, an honorary member of the Hungarian and German Societies of Psychology, and of the International Association of Applied Psychology. He also took an active part in many national and international congresses, those held in North America, including the Congresses at New Haven (1929) and Montreal (1954).

In addition, he was co-editor of various Italian psychological reviews and contributed to some non-Italian periodicals, including *Psychological Abstracts* and, more recently, *Psychological Reports.*

Ponzo left about 280 publications. He busied himself with both general and applied psychology. Indeed he always maintained the essential unity of psychology, that pure and applied research complete one another and are interdependent. In everything he did, he exhibited intense activity as scholar, research-worker, and organizer.

Let us examine some of the subjects in general and experimental psychology, in which Ponzo specially developed his dynamic teleologic functional principles.

Sensory and perceptual processes. A group of very early researches, encouraged by

Kiesow, are of a histological nature and concern the distribution of the taste buds in the pharynx from the choanae to the upper esophagus. Later this led to his psychophysiological interpretation of the act of swallowing as an example of a purposive perceptive-motor synthesis. As further examples of this synthesis, Ponzo examined some after-effects produced in the blind's tactile-kinetic perception of rigid forms. In other researches he pointed out the errors, and some illusory phenomena accompanying them, that occur when the subject is asked to localize the point of application of the stimuli of touch, pain, and heat, especially when some cutaneous areas stimulated are artificially displaced from their normal site. For this work he devised two pieces of apparatus: a dermolocalimeter and a special esthesiometer.

He also took up, at various times, intuitive appreciation of the number of elements united in figural complexes. He paid special attention to the changes from the sensations of weight to those of lightness, and to various concomitant illusions, regarding either the stimulus or the body, when the stimulus—always much greater than threshold—is made to decrease continuously. He found that the passage from the perception of weight to that of lightness occurs through a moment of perceptive annulment of the stimulus. Thus, he considered reports of weight and lightness to be expressions of the same intensive and qualitative continuum. Also, in the field of sensory-perception, he worked on after images, the action of some local anesthetics and drugs; and the fluctuation of perceptive attention and of attention-span.

In the field of *imaginative and representative processes*, Ponzo studied "collective images," as he called them, that are closely allied but not equivalent to stereotypic images, and the *déja-vu*. He also promoted researches into the orientation of remote space.

Psychophysiological study of breathing. He devoted himself to the study of some phenomena of induction (consequent on the perception of external rythms) of respiratory frequency, the pneumographical analyses of the associated processes of the recognition and the naming of objects, the pneumographical signs of inhibition (both unconscious and voluntary) of verbal associative responses, the familiar pneumographical affinity during expressive speech, the analyses of pneumographical modifications during various forms of perceptive attention and of silent mental activity, accompanied or not by interior formulation of speech. In these researches he sought to demonstrate that breathing should be considered not so much as a symptom or a concomitant of various psychic activities, but rather as psychophysical behavior itself, as adaptation, and also as an expression of the unity of the person in action.

Psychology of action and of the comprehension of the personality. The dynamic and teleologic path followed by Ponzo was developed by him above all when he occupied the chair of psychology at Rome University. His first lecture (1931) was entitled *Present-day trends of experimental psychology as the science of the dynamism of psychic life*, and by the term, "psychology of acting" or "psychology of action," he upheld his theoretical point of view in psychology. He argued that every action with its particular end is the expression of the whole personality, so that the single action may be chosen as the ultimate indivisible element, as the concrete functional unity from which psychology can start the study and comprehension of

the personality itself. He insisted, however, that the action should be observed in its entirety, not only in its external manifestations but also in its "internal elaboration." The German Society of Psychology, in graceful recognition of the importance of Ponzo's work on both external and internal actions, made him an honorary member on his 70th birthday.

Psychomotor processes. In the above theoretical perspective of the psychology of action, Ponzo gave an importance to motor processes which must be considered both as an expression of the integration of mental and organic activity in the indivisible unity of the person and as a vehicle of action (psychomotricity). Thus the research undertaken by him and his assistants on "recovery time" (with its adaptive function in simple reaction times in series), the analyses of the phases of single serial acts of tapping, group-motor operations, the motor reactions emotionally disturbed by unexpected stimuli (analysed by photocyclographic technique), correspond to his criteria regarding the methods for studying voluntary motor processes. "These methods," he wrote, "starting from the *external form* of the motor action, should aim at making clear all that happens internally as dynamic impulses, contrasts, inhibitions and choice of voluntary action." He also promoted, in the Institute he directed, research into instinctive animal behavior.

Ponzo also devoted himself assiduously to *applied psychology.* He began by interesting himself, when an alienist in Turin, in finding work for mental patients discharged from the hospital. Thereafter he studied mechanical workshops in a professional school examining individual differences in vocational aptitudes and in the training of vocational capacities. In Rome, he threw himself into educational and vocational guidance (seeking above all to establish a national standard as a safeguard against dangerous individual initiatives). He concerned himself with the study of some specific occupational profiles, with the scientific management of labor, with the employment of disabled ex-servicemen, problems of didactics, and with adult education. More recently he took up the psychology of motion-pictures and social work. For the latter, he founded a school—now in its fifteenth year—at the Institute of Psychology, University of Rome.

In his manifold activities, his fighting spirit never flagged. It was innate through the example and traditions of his family and was proved both in the First World War (for which he volunteered) and no less in his long, arduous campaign for the recognition of psychology in Italy (almost incredible to many non-Italian psychologists), a campaign in which he was preceded by Kiesow and De Sanctis and then supported by Gemelli and by others. Into this effort Ponzo threw all the ardent and sometimes impulsive dynamism, which undoubtedly motivated his "psychology of action."

All the sadder, in view of his enthusiasm and vigor, was the complete inactivity forced upon him at the end by the relentless progress of his malady.

University of Rome LEANDRO CANESTRELLI

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Measurement: Definitions and Theories. Edited by C. WEST CHURCHMAN and PHILBURN RATOOSH. New York, John Wiley and Sons, Inc., 1959. Pp. viii, 274. $7.95.

"The trouble with the idea of measurement is its seeming clarity, its obviousness, its implicit claim to finality in any investigative discourse" (p. 163). This volume contains an antidote to much of the nonsense written for and by psychologists on the topic of the theory of measurement. Possibly, no topic in the philosophy of science is more relevant to the needs of psychologists, and few topics have had so much written about them that is so confusing and discouraging. On the one hand, as Stevens comments regarding N. R. Campbell's "classical" position, "Why, he might have asked, does the psychologist not give up and go quietly off to limbo?" (p. 23). On the other hand, we have heard a most distinguished and revered Nobel Laureate, whose intellectual achievements in atomic physics verge on the incredible, offer to share with us, as a rationalization for our own lack of impressive achievements in psychology and the social sciences, his principle of complementarity, a principle quite appropriate to canonical conjugates, but a philosophical guide which is called, quite properly, a "doctrine of despair."

The reviewer believes that psychology has much to learn from the structure and problems of physical science, but that most of us cannot hope to master physics, hence, when we use analogies from physics, it especially behooves us to try to get our facts right at the start and to reflect earnestly upon those facts as we go along. For brevity, clarity, and authoritative accuracy, the reviewer believes that the chapters by Caws, Margenau, Pap, and McKnight can scarcely be excelled. No one who reads them with care will continue to regard the uncertainty effect as a consequence of the interference of the measuring operation with the measured object, or be tempted to imply that our troubles, great as they are, are caused by such ultimate limitations. Indeed, some may take heart from the views of Born and others (p. 200) that the concept of complementarity is an outmoded crutch even for physics, and know that we in psychology do not yet and may not ever hurt in such a way as to need it. In short, we can face the future with Einstein's courage, as "half-tamed metaphysicians."

On the meaning of measurement, the chapters by Caws, Stevens, and Churchman take somewhat different approaches, but where they intersect they are remarkably harmonious. In view of the controversies which have beset this topic, such harmony is impressive, and I believe that it represents real clarification of and agreement on issues rather than prejudicial selections among the authors. If this is true, the timing of the symposium and book, and the selection of participants reflects admirably on the editors. Caws points out that, "The true function of measurement is to link mathematics and physics" or science (p. 13). It is perfectly clear that no science can be reduced to mathematics, but the hope seems to be entertained never-

theless that science can be reduced to logic, which is exactly what cannot be done. The mathematical nature of a science is forever analogical. As Menger shows (Chapter 5) the variables of a science cannot be regarded as a class of numbers but must be described as a class of ordered pairs, one member an element from a class of entities from the domain of the science, the other member a number, selected from the class of numbers by a principle from the domain of the science. Caws goes on to point out that "the construction of the pair in this way is a matter of convenience" (p. 11). Accepting all this, Churchman goes on to investigate "convenience." He sees that the consequence of this approach is that measurement may be regarded as a decision-making activity. As he explores the implications of this view, he seeks no final decision-functions, but he proposes some of the more intriguing ideas in the book. He mentions the "costs" and "returns" associated with decisions as to choice of language, specification of objects and environments to which his measurements shall apply, standardization of results, and evaluation of their accuracy. For example, in choosing language, the more precise it is the less widely will it be understood, etc. If measurement does indeed have only a pragmatic basis, as seems likely, then the decisions as to what shall be done and how will involve further development of the ideas Churchman briefly outlines. He goes on to remark that scientists might as well get about the task, for although science may be immortal, all existing decision models assume penalties for delay!

Part II, "Some theories of measurement," contains three rather technical papers by Karl Menger, Patrick Suppes, and R. Duncan Luce. They deal, respectively, with the need for explicit separation of the mathematical connectives from the scientific objects they connect, with mathematical versus empirical meaningfulness and the necessity in the latter case of a third truth value, meaninglessness, and with the demonstration that both utility and subjective probability form Fechnerian sensation scales if statistical independence of preference and expectation may be assumed. That this independence is a necessary condition for these scales is not proved but seems likely. Since human observers do not seem to expect and prefer independently perhaps subjective probability and utility, in Luce's sense, cannot be made to yield such scales.

In the final part of the book are presented two experimental applications of measurement and decision theories in the social sciences. Coombs shows that the traditional measures of inconsistency of preferences are not enough to scale an ordered set of objects when one of the central objects represents a subjective ideal. Here it is shown that pairs from opposite sides of the ideal give inconsistencies of preference of a different order of magnitude from that found when both members of the pair are on the same side of the ideal. In the last chapter, Davidson and Marschak find that a stochastic theory of decision is superior to actuarial values of wagers or the degree of risk in predicting choices among wagers. This should not be news.

Although there is inevitable variation in difficulty owing to the highly technical character of some of the papers, the exposition is quite good throughout the book, and two of the chapters, Margenau's and Stevens' struck the reviewer as being examples of exceptionally elegant prose. Most psychologists who read the book will be familiar with Stevens' position and will know that he has had a difficult time gaining adherents outside of psychology. His chapter again finds him patiently and earnestly trying to meet the insuperable objections to his conceptions of the bases

of measurement while continuing to provide a wealth of data which prove that one can make distinguished contributions without convincing the majority that one's rationale is correct.

In summary, the reviewer recommends this book most highly. Short as it is, its study will be richly rewarding and encouraging to those who face the bleak fact that making science is itself one of the great acts of faith.

The University of Wisconsin DAVID A. GRANT

Personality. By J. P. GUILFORD. New York, McGraw-Hill, 1959. Pp. xiii, 526. $7.50.

Guilford has written a textbook in the tradition of Woodworth's *Experimental Psychology* in extent of coverage, but narrower in breadth of point of view. There are over 900 references. Most of the major research in the field, and much of the minor, seems to have been covered.

The book can be divided into three parts: basic concepts and approaches comprising the first five chapters; a review of methods of measurement in the next eight chapters; and a dimensional analysis based on available research in the last five chapters. The summaries at the end of each chapter are splendid integrations of the mass of material covered and should provide excellent study guides for students. There is a short appendix consisting of an introduction to the basic idea of factor analysis. The name and subject indices seem thorough.

Personality continues to be a difficult domain to delimit, yet sufficiently enticing to the psychologist concerned with human behavior that he would not think of abandoning it for simpler fields. Guilford grapples courageously and in deceptively simple language with the basic problems of definition and approach in the first five chapters. The discussion is always straightforward. Many teachers will want to refer their students to these chapters even if they prefer a different organization of the area and an emphasis on other variables. Particularly helpful for the professionally naïve student will be the sections dealing with the varying goals of divers professional groups in their respective studies of personality, and the clarification of the aims of the scientist. Nevertheless, the current state of the field is reflected in that definitions continue to be imprecise and aims not always clear. To take a very simple example, a trait is defined as "any relatively enduring way in which a person differs from others" (p. 8-9). Surely "any" is not to be taken literally (*e.g.* we would hardly want to regard nose lengths as personality traits), but if not, what are the limits of our study? Another ambiguity is hidden in "relatively enduring": relative to what? Just what is the order of time in which we are interested? These questions may be quibbles from the practical point of view, but theoretical discussions in personality often seem to stem from unspoken implicit assumptions about the very limits of the field. At the same time, it should be pointed out that Guilford's attempts not to imbed his definitional statements in masses of verbiage make his discussions clearer than the great majority of such essays. The ambiguities are inevitable in a growing field, and all we can require are continuous refinements based on further research and improved conceptualizations.

The picture is very different in the last thirteen chapters. Here Guilford's biases have a decidedly narrowing effect on the kinds of data he considers, and result in a peculiarly static, descriptive organization of research finding—taxonomy carried

to an extreme, without the redeeming feature of the kind of hierarchical structure that could lead to the hypothetical formulation of functional relations for research purposes. All that seems to be left is to seek endless new dimensions through factor analysis. Presumably, more and more detailed compendia of the results of such analysis will appear in the future. It is certainly true that this book contains the most thorough summary available to this point of the important research which uses the correlation coefficient as the basic statistic. One searches in vain, however, for a statement concerning a *chain* of behavior, and the necessary and sufficient conditions for its occurrence. The frustration-aggression hypothesis is mentioned (p. 447), for example, but there is no reference to the original monograph, to the studies of Child and Waterhouse nor to the Lewinian work. The work on regression is not mentioned at all.

Generally, there seems to be reluctance to refer to overt behavior as the final testing ground for the validity of proposed dimensions. In this respect, we see revealed the ultimate weakness of the basic data on which many factor analyses are based. As has been pointed out so often before, no amount of statistical manipulation can convert a response (*e.g.* to a questionary), the determinants of which are unknown, into part of a meaningful factor.

There is reliance on factor analysis as the sole method of organizing and interpreting research data in this book. There is insufficient evaluation of the worth of the raw data and of the relevance of the final factor matrix to behavior. Instead, Guilford has frequent recourse to the notion of "factorial validity," defined as the validity of a trait found by factor analysis. "The degree of factorial validity of a test is indicated by its factor loading" (p. 112). One wonders if the circle described by this perpetual circularity of the factor analytic methods is large enough to be useful for very much longer.

To sum up, this is among the very best textbooks based on factor analytic studies of personality yet produced. The over-all orientation is probably too narrow for most introductory courses in personality. The book should prove very useful, however, as auxiliary to a more broadly-gauged text, particularly the first five chapters, but also, for discussion of selected samples of the factor analytic literature. Every psychologist can learn something from this book, and those in the area of personality will certainly want to be familiar with it. The final picture of personality offered seems thin and static, but reflects more rigorous thinking than is usually found in books on personality. There is little discussion of serial behavior linked by functional relations.

University of Pennsylvania JULIUS WISHNER

Anxiety in Elementary School Children. By SEYMOUR B. SARASON, FREDERICK F. LIGHTHALL, KENNETH S. DAVIDSON, RICHARD R. WAITE, and BRITTON K. RUEBUSH. New York, John Wiley and Sons, Inc., 1960. Pp. viii, 351. $7.75.

This book reports a six-year program of research on anxiety in elementary school children, a program carried out by the principal author and his graduate students and colleagues. A synthesis of the published studies is presented and a number of unpublished investigations are reported. The objectives were practical; namely, to understand the correlates of test-anxiety in elementary school children and to de-

velop measures which would have diagnostic and prognostic utility in the school. The book reports more than the development of procedures for measuring test-anxiety and general anxiety in the classroom. The investigators have brought to bear on this problem both clinical sophistication and familiarity with psychoanalytic theory as well as ability to use observational, experimental, and psychometric techniques. For this reason the treatment of the area is superior to many research projects which merely involve administration of a test and quick return to the laboratory to add up the scores. The children were observed, their parents were interviewed, projective techniques were used, special experimental techniques were devised to test promising leads, and both interviews and tests were subjected to critical analyses for the possible contributions of response sets or defensive reactions. The investigators stayed with their original problem until its major parameters were outlined. A number of possible contributing factors were tracked down such as social class, illnesses, and parent-child separations.

The major findings were as follows: The orally administered group questionary of test-specific anxiety correlated, though unimpressively, with teachers' ratings of anxiety, showed higher anxiety scores for English school children subject to greater test-pressure, and showed a negative relationship with *IQ*; *S*s with high test-anxiety scores S_{HA} (*HA-S*s) showed higher *F*-scores and less use of color and shading on the Rorschach, they showed more mutilation and rigidity, less smiling and playfulness in human figure drawings, and they learned less in a learning-situation where time and competition were emphasized; *HA-S*s did better but showed more caution than *S*s with low test-anxiety scores (*LA-S*s)—at certain intelligence levels—on the Witkin Embedded Figures Test in which time is not emphasized; *HA-S*s scored higher than *LA-S*s on the Porteus Maze Test which minimizes time pressure but lower than the *LA-S*s on the Stroop Maze Test and Iowa Basic Skills Test, both of which involve more time-pressure; mothers of the children studied showed more defensiveness than fathers in assessing their children; only information obtained from the fathers discriminated *HA-S*s from *LA-S*s adequately; differences between *HA*- and *LA*- boys were more striking than between girls, and there was evidence that *HA*- girls were not comparable to *HA*- boys in several ways. Contrast between *HA-S*s and *LA-S*s used individuals or groups matched on *IQ* and other relevant variables.

The authors attribute the development of test-anxiety to an inordinate concern of the parent with evaluation and social acceptability. Presumably the parents ignored the child's capabilities and age-specific needs. The result in the child was a derogatory self-image, unduly inhibited direct expression of aggression, and extreme dependency on others for approval. It is this last characteristic which was thought to contribute to the child's special difficulty with tests of the type which create the impression that *S* runs a high risk of not living up to expectations. This risk presumably threatens dependency needs.

One shortcoming of the research is the tendency to view the development of test-anxiety as entirely a result of the parents' contribution, the parent being seen essentially as invariant. The possibility is not considered that some congenital factor in the child could have contributed to the development of test anxiety and that the parents' behavior might be a reaction to the coping difficulty of the child. Search for and study of *HA*- and *LA*-children in the same family should be given early

attention to achieve more precise understanding of determining factors. It has not been difficult in other areas of psychopathology to locate children showing and not showing a symptom pattern in the same family. Very different patterns of behavior have been found in mothers toward different members of identical twin pairs, one of whom has a congenital disorder.

Another shortcoming of the book is a tendency to reject findings which do not fit in with the theory, citing much higher research standards than are applied to findings which do fit the theory. For example, the authors posit a relationship between school-phobia and test-anxiety. Accordingly, relationships between school-phobia and *IQ* should be negative to fit in with the authors' findings. Studies reporting strong positive relations are dismissed with the statement that replication is needed. Statistical characteristics of teachers' ratings of test-anxiety are criticized sharply when they fail to correlate substantially with the questionary developed, and yet the discussion ignores the fact that these ratings show they are capable of correlating at a quite satisfactory level with objective measures of achievement. These shortcomings are balanced by the authors' evident willingness to admit and to describe in detail a large number of failures to confirm specific hypotheses. A path of trial-and-error in improving the conceptual network with successive studies is clearly outlined.

Biosocial Growth Center, CI, NIMH — Richard Q. Bell
National Institutes of Health

Art and Illusion: A Study in the Psychology of Pictorial Representation. By E. H. Gombrich. New York, Pantheon Press, 1960. Pp. xxxiii, 466. $10.00.

Nothing could be more familiar than pictures, and yet nothing is more strange. Almost any statement about a picture is controversial. If one says, for example, that a painting or a photograph yields an "illusion of reality," all sorts of formidable questions arise. Here is a book which states the questions even if it does not fully answer them.

Representation, like speech, is one of the marks that distinguish the human animal. Pictures have been made by men since long before the dawn of history, and they have probably been puzzled over and talked about by would-be critics for just as long. The art of painting has a recorded history of debate which began with the Egyptians and the Greeks. Plato disapproved of painting because it was deceptive. Later on, paintings were admired because they "mirrored nature." Today, in the photographic era, it is argued that painting should not be representational, for the duty of the artist is not to create a likeness but rather to teach us to see in new ways. The discussion of these questions is always interesting but often unintelligible. This book, in contrast to most, is highly readable.

The reason for all the confusion about pictorial representation over the centuries is that it is part of a deeper confusion about the act of vision itself. We do not understand the making and viewing of works of art because we do not understand perception. The author of this book realizes that theories of art imply theories of perception, and he has done something unusual; he has read very widely and sympathetically in the experimental psychology of perception.

Gombrich is an art historian, in the European tradition of general and far-ranging scholarship. He knows the paintings of the ages and he has put some 300 illustra-

tions into his book, but, along with the fine art, he has included caricatures, posters, sketches, illusions, and diagrams. Each has its point. What he wishes to demonstrate is that every picture is not only an act of perception but also, in a sense, a psychological experiment. The experiments of the picture-makers are unformulated and uncontrolled, and they are not very satisfactory to the scientist. The artist, however, is always an investigator, and there is usually a hypothesis which he wishes to test. The history of these hypotheses is not part of the history of formal knowledge, but a humanist who also sympathesizes with the scientific tradition can go far to make them explicit.

The author has searched the whole of modern psychology for guidance, with complete willingness to learn. He has none of the rebellious anti-scientism which sometimes infects the critics of art. He has borrowed from the ideas of many theorists, using them when he can.

It would be satisfying to be able to report that this excursion into psychology has solved the problems of visual art, but it has not. The psychologies of perception embody too many deep contradictions for that. The author's eclectic approach results in a considerable clarification of certain problems of pictorial perception and the exposing of a good many fallacies but not a systematic theory. The formulas that he finds helpful in one chapter are not always consistent with the hypotheses he finds illuminating in the next.

This book should be at least as useful to psychologists as to artists. The Gestalt theorists, particularly Koffka, made us face up to the subtleties of visual phenomena. The determinants of figure-on-ground and of the organization of the figure are now listed in textbooks, but the actual scope and variety of the visual information which can come from tracings on a surface is barely suggested by these so-called 'laws.' Gombrich illustrates many of the higher orders of visual stimulation on nearly every page of his book. The discoveries of painters have been far more elaborate than the discoveries of psychologists, if less rational, and Gombrich shows that they are at least potentially investigable. The student of perception is tempted to limit his research to what he can experimentally control by the methods he has been taught. This book will widen his horizon and stimulate his ambition.

Cornell University JAMES J. GIBSON

Science, Medicine, and Morals: A Survey and a Suggestion. By CHARLES E. RAVEN. New York, Harper and Brothers, 1959. Pp. 189. $3.50.

This interesting book should appeal to those who hold that facts have implications, *i.e.* those who reject the phenomenological tradition and those who reject the so-called Cartesian dichotomy of body-mind. It is a straightforward, logically organized work of general interest, although it is particularly an admonition to the physician to expand his historic role.

Dr. Raven has been trained in the 'hard sciences' and is a theologian. He believes that the present is a period of crisis for civilization, as man has never before held within his grasp the power either to destroy himself or to inaugurate a golden age. His book is an attempt to explain the tendency for East and West to become more and more disoriented and estranged from our cultural heritage in terms of the impact of the scientific method.

Unfortunately science and religion have grown to distrust each other and to feel

that a basic conflict is unresolved between them. The historical mechanisms resulting in this dichotomy are discussed: science and religion both have been at fault; science has felt it necessary to insulate its field from philosophy, and faith and religion; while religion (Christianity) has reverted to the pre-Christian view that God is a great engineer or watchmaker whose function is to create a perfect machine. Raven emphasizes that fatherhood is the only synonym for creativity which a Christian can legitimately hold.

Raven sets down the points on which science and religion can agree, *i.e.* that the universe is a cosmos, an order, not random, promiscuous, unpredictable; that it hangs together (although he does not straightforwardly advocate monism); that life is teleologically oriented; that suffering and death are essential to the teleological development of human life; that in this teleological process of life we must realize and must accept the idea that the novel and the new can emerge. Raven also believes that the new physics of Einstein and Heisenberg and that the Gestalt point of view in psychology serve to reunite the sciences with a theology which is recovering its legitimate point of view. He thinks that the new science and theology agree that novelty and creativity exist in the universe.

Raven's thesis is to trace the history of the relationships between religion and science from the point of view of medicine. Concerned not only with the healing of disease but with the promotion of health, the medical man, by his training and experience, is equipped to treat humanity neither as a machine nor as a ghost, neither as the robot of the biologist nor as the fallen angel of the theologian, but as a total human being. The physician must get rid of the artificial barrier between the body and the spirit.

Raven believes that the medical man, throughout the history of science, has served, by his insistence that the body and the spirit cannot be separated, to keep a sense of proportion when the theologian floated off into a world of dogma, and the scientist ran after partial and diversive speculation. He thinks that medicine's insistence on treating the whole being, its view that health is not simply the alleviation of disease, its idea that the person and his environment must be involved, can save us. Medicine must evolve and it is evolving a new emphasis, not upon the particular and sectional treatment of disease, but upon the maintaining of health, the building of an environment which will promote health, and the encouraging of activities which make for a healthy mind in a healthy body and a healthy individual in a healthy community. Medicine still needs the man of scope as well as the specialist. Thus, Raven concludes, that the new physician must study the ills of humanity whether they be organic or social.

University of Texas ROY A. JOHNSON, JR.

The Child's World: His Social Perception. By FRANK J. ESTVAN and ELIZABETH W. ESTVAN. New York, G. P. Putnam's Sons, 1959. Pp. xiii, 302. $4.95.

The Estvans' book reports the results of research designed to discover "what each child's world is." More specifically, the research dealt with "the relationship between elementary-school children's social perception and differences in types of community lived in (rural-urban), sex, grade (first-sixth), and intelligence."

The problems to which the authors have addressed themselves are of utmost theoretical and practical importance. Unfortunately, however, their approach to them

seems inadequate largely because they used only a single, and unfortunately highly artificial, technique. It consisted of 14 pictures of life-situations depicting rural and urban and upper and lower class scenes and adult-child interactions. These were shown, in individual interviews, to 88 children who were asked "What story does this picture tell?" Their responses were evaluated from the standpoints of *recognition* ("what the scene represented to the child"); *perceptual field* ("the where or when orientation ascribed to the scene"); *perceptual attitude* ("how the child responded to the situation").

The research suffers from several important deficiencies. For one thing, not all the subjects reacted to the same or, from the investigators' point of view, to the most critical aspects of the stimulus-pictures. Consequently, the pictures often failed to elicit the attitudes for which they were designed. For example, a picture of a dam that was supposed to tap attitudes toward conservation and a school scene that purported to "elicit ideas about education and human relations experienced by children in school" frequently evoked statements that were irrelevant to these central social themes. Even, however, when relevant attitudes (*i.e.* those related to what the investigators considered the most significant aspects of the pictures) were expressed, there is little reason to believe that these accurately reflect the child's way of perceiving, or reacting to, real-life situations. Furthermore, the number and kinds of situations sampled were limited; in spite of the title of the book, they are hardly representative of "the child's world."

In this reviewer's opinion, the cardinal defect of the work involves an omission —or at least a missed opportunity. The application of some form of Piaget's *methode clinique* during the interview might have yielded invaluable information about how and why children perceive and structure their world as they do. Instead, only direct, simple, and immediate responses were elicited and, partially bcause of this, the analyses were gross and superficial. As a result, the investigators learned much less than they could have about the conceptualizations underlying the children's reactions.

Despite these limitations in procedure and analysis, some interesting findings related to sex- and age-differences in ways of reacting to social situations were discovered. In view of the small number of subjects and the frankly exploratory nature of the study, these must be considered only suggestive, although the authors sometimes generalize their findings rather broadly. For example, they summarize some of their findings with the statement that "the evidence is conclusive, therefore, that no real trend in generalizing about types of life situations develops during the elementary school years" (p. 258).

If the findings reported in the book are reaffirmed in more extensive and intensive systematic research, they may prove useful not only in educational planning but also in the formulation of broader theories of social perception or in the modification or extension of existing perceptual theory. Unfortunately, however, the authors offer only very general hypotheses—usually referring to "learning and experience" —to explain their results. No attempt is made to integrate the findings of the present study with the existing body of facts or theories of perception.

University of California, Berkeley PAUL H. MUSSEN

Trends in Content Analysis. Edited by ITHIEL DE SOLA POOL. Urbana, University of Illinois Press, 1959. Pp. 244. $7.50.

This symposium is based on papers prepared for a conference sponsored in 1955 by the Committee on Linguistics and Psychology of the Social Science Research Council. The sponsorship is reflected in psychological and linguistic emphases in which the present book differs notably from previous major publications on content analysis by Lasswell, Leites, *et al.* (*Language of Politics,* 1949) and by Berelson (*Content Analysis in Communication Research,* 1952.) The marriage between psychology and linguistics, however, hardly comes to fruition: the reader is left with a heterogeneous collection of papers not really addressed to a common audience.

Psychologists will be interested principally in major contributions by Charles Osgood and by George Mahl, and in the editor's brave attempt at integrative summary. Osgood discusses the assumptions and applications of three techniques: his own developments of evaluative assertion-analysis and contingency content-analysis, and Taylor's 'cloze' technique. The initial part of Mahl's paper is a closely reasoned analysis of verbal productions as instrumental behavior, in which he takes issue with the 'representational model' implied when (as with Osgood's evaluative assertion-analysis) lexical meanings are coded at face value. Inference from message content to communicator's intent involves pragmatics rather than semantics, implies Mahl, and reinforcement theory provides the key when enough of the context of communication can be observed. In the empirical part of his paper, dealing with his research on indicators of anxiety in psychotherapy, Mahl, however, fixes on expressive features of language behavior that are neither instrumental nor representational in Osgood's sense.

Other contributions to the book include a defense of qualitative, nonfrequency methods in the analysis of wartime propaganda (Alexander George), two chapters mainly about problems of folkloristics (Saporta & Sebeok; Armstrong), and a relatively uncritical examination of the potentialities of content analysis for biography and history (Garraty.)

As Ithiel de Sola Pool notes in the concluding chapter, the papers primarily sample only one segment of possible interests in content analysis, that concerned with inferences about the communicator. The heterogeneity of conception and approach even here raises questions about the meaningfulness of content analysis as a category of research technique. The early enthusiasm in communication research for the narrower conception of content analysis reflected in the Lasswell and Berelson books having subsided, does the present book, edited by a former associate of Lasswell, represent a renaissance of interest? Not, thinks the reviewer, in a subdiscipline or technology as such. Just as the enchantment of discovering that we are talking prose fades, so the research necessity for analyzing prose in many ways for many purposes may provide grounds for a profitable conference but hardly sustains the enthusiasms of a specialty. The heterogeneity of the present book, its very failure to establish substantial common ground among content analysts, may be a sign of maturing research interests.

Of the trends suggested by the book, the most promising and the one on which the papers show most genuine convergence is toward techniques that tally not simple frequencies but the contingent association of content-categories. Developments since the conference support the view that contingency analysis is a powerful analytic tool

Were the papers written today rather than in 1955, the conference would undoubtedly have noted how the availability of high speed computers has immensely extended the flexibility and range of this approach.

University of California, Berkeley M. BREWSTER SMITH

Sampling Opinions: An Analysis of Survey Procedure. By FREDERICK F. STEPHAN and PHILIP J. MCCARTHY. New York, John Wiley & Sons, Inc., 1958. Pp. xxi, 451. $12.00.

This book is intended to fill a niche between the many mathematical treatments of sampling theory and problems and the perhaps even greater number of treatises on methods of measuring human opinions, attitudes and wants. Because the authors feel "that preoccupation with mathematical models now feasible will detract attention from many important problems not yet incorporated effectively in these models," they have undertaken to describe in almost entirely non-mathematical terms, the relationships between sampling procedures as usually conceived, and all the parts of the entire survey procedure.

An understanding of the term sampling in the at once restrictive and highly general senses employed by the authors gives meaning to the organization and content of this book. Discussion is restricted to only such surveys as are concerned with opinions, attitudes, and wants; thus the concern of a survey is with *a universe of human responses.* In a general sense then, a survey may be thought of, in total, as a sampling device to obtain a sample of measurements of responses from a universe of responses; those sample data to be used in deriving findings and conclusions. This brings squarely within the purview of sampling every aspect of the survey process from problem formulation to final data, including the interdependent relationships of all the parts.

The first section of the book introduces the reader to uses, procedures, problems and terminology of survey operations. Through the use of some simple descriptive models, essential aspects of a variety of sampling operations are characterized and discussed in relation to the survey processes within which they are embedded.

The second part examines, at a level which requires from the reader at least some degree of statistical sophistication, a number of empirical terms, error and sources of error for a wide variety of sampling procedures. It would seem that the present day role of quota sampling hardly justifies the amount of attention the authors give it. However, the presentation makes clear the difficulty and necessary complexity of such evaluation, and emphasizes the danger of doctrinaire conclusions regarding the value of any one sampling procedure.

Part three discusses the practical problems of designing a survey and putting it into operation. The presentation is realistic and reflects a wealth of experience in all aspects of survey research. It beautifully points up the fact that every survey must be a "best" compromise which tries to balance costs, respondent accessibility, sample size, clustering, accuracy checks, supervision, training, time and many other factors. It is a reminder of the host of practical problems faced by the practitioner, of the as yet limited role in the total process of mathematical models, of the degree to which we must depend on experienced judgment.

Despite the fact that this book is uneven in the level of statistical and survey sophistication demanded of the reader, leaving it open to the charge that it misses

both the expert and the non-expert, reading it probably will benefit both because this is an eminently *sensible* book which provides excellent perspective on sample surveys. As a matter of fact, probably any survey or any survey organization would benefit by checking operations against the procedural generalizations given in the last section of the book.

For many people, reading this book may be more task than recreation, but anyone, with a need to understand or evaluate surveys of opinions, attitudes and wants, should profit from doing so.

The Brookings Institution FRANKLIN P. KILPATRICK

Leadership, Psychology, and Organizational Behavior. By BERNARD M. BASS. New York, Harper and Brothers, 1960. Pp. xii, 548. $6.50.

This text on the social psychology of leadership is a comprehensive, lucid, and faithful account. The bibliography contains 1155 references and the reader comes quickly to appreciate the care with which these materials have been culled and cross-classified to make available the ideas and findings they contain. In part because of the author's care, and, in part, because Bass writes with straight-forward clarity, one has the sense of being well informed without becoming lost in detail.

The material is organized according to those problems most salient for the selection, development, and evaluation of persons who lead. After some preliminary chapters which sketch the status of early knowledge on the topic and which define key terms, we are introduced to methods for measuring and evaluating leadership. Then follow substantive chapters on the relations between leadership and group effectiveness, varieties of leadership-style, the development and generality of leadership-qualities, the social and personal sources of a leader's influence, and strategies for dealing with various conditions which debilitate a group's cohesion. Data from psychology, history, and the several social sciences are brought to bear on these problems.

I have stressed the character of the problems around which the book is organized, because these seem to provide a major clue for an evaluation of Bass' work. We may contrast his perspective, so heavily determined by the psychology of assessment, with quite different bases for organizing these same materials.

All recent works on leadership employ the notion that it is a social relation rather than a property of individuals. Thus Bass says that leaders are persons who direct their behavior "toward the goals of changing another's behavior (p. 23)." This conception proved marvelously productive in guiding empirical investigations. The conception is so broad, however, that very few social roles can be completely excluded from its compass. To take an extreme case, the slave who works diligently in order to mitigate his master's wrath is, by this conception, a leader.

To avoid having to treat all interaction as an instance of leadership, one may employ various strategies. The more sophisticated works on public administration and business management begin with descriptions of the structure of particular types of organizations, usually large bureaucracies, and proceed to state the conditions under which persons can be influential on the broad policies of such enterprises. The characteristics of the type of organization and of its component tasks and roles provide a coherent framework for these treatments, including the consideration of those characteristics of individuals which may facilitate or impede their exercise of major

influence. The same approach is used by a number of students of leadership in small groups—Bales and Festinger, for example.

The major limitation of Bass' work is that he lacks such a coherent framework. If leadership is almost coterminous with all social roles, the nature of the qualities required in leaders and the impact they have on a group's effectiveness vary as do the roles. Since the varieties of role are legion, so also are the varieties of leadership. Although Bass' chapter headings give one a sense of order and coherence, this is the order of lexicography; not of functional relations in nature. As a result, Bass' great objective of subsuming his myriad of references in a parsimonious and coherent framework goes unrealized. It appears that, if one is to employ a social conception of leadership, he must also employ a social-structural framework for organizing the phenomena of leadership.

The University of Michigan GUY E. SWANSON

Tropical Childhood: Cultural Transmission and Learning in a Rural Puerto Rican Village. By DAVID LANDY. Chapel Hill, The University of North Carolina Press, 1959. Pp. xii, 291. $6.00.

This book is the author's account of eight months of research with 18 children in 18 separate lower-class families living in a small relatively isolated community, whose chief economic activity is the raising of sugar cane. Landy calls this community "Valle Card."

In the first chapter Landy presents a careful description of his methodology. In addition to the usual ethnographic methods, the author gathered systematic data from structured interviews, modeled after the one used by Sears, Maccoby, and Levin (*Patterns of Child Rearing*), with the parents and with children playing with dolls. Considerable care was taken to obtain inter-rater reliability for the ratings of the interview-data. The author explains that he has not included photographs in his text. He excluded them in favor of tables of interview-ratings which he presents throughout the chapters on child training. In view of the smallness of the sample, these tables sometimes appear rather pretentious. Since the size of the samples does not permit tests of the influence of demographic variables on the interview responses, the tables are, for the most part, simply distributions of a particular rating. Furthermore, the meticulousness of this tabular presentation is marred by the author's inadequate description of his rating-procedure.

Following the chapter on methodology, the author presents a detailed, if somewhat repetitious description of the family life of this community. After a general description of the village in terms of geography, living conditions, religion, economics, class composition, and the political and kinship structure, Landy describes the status and role-patterns of his sample families, parental attitudes towards marriage and children and social practices.

Although the presentation is sufficiently complete that the persistent reader may obtain a fairly good picture, the organization of the book is poor. The author divides each of his main topics, culture and society, bringing up the child and coming of age into two chapters—I and II. The basis for this chapter division is obscure. The organization is neither chronological nor topical. After a discussion of several areas of children's behavior and parental treatment of them, the author brings his reader abruptly back to dscription in the subsequent chapter. Almost all topics on socializa-

tion—weaning, toilet training, aggression and dependency, and disciplinary techniques—are spread over two or more chapters. This leads to unnecessary diffuseness and repetition.

Following the description of life in Valle Card, there are two chapters comparing the Puerto Rican sample with the New England sample studied by Sears, Maccoby, and Levin. Since the interviews are much the same the comparison is extensive but, lacking cultural interpretation of differences, somewhat meaningless.

As a whole this is a useful source book for anyone interested in cross-cultural studies of children. It represents a step forward from the traditional methods of anthropology. The book is, however, rambling and lacks focus. A number of inconsistencies in parental discipline and expectations for their children are pointed out but never fully discussed. In his concern for careful description of variations within samples, Landy has lost much of the sharpness of the general enthnological description. Whether his gains in accuracy of reporting are sufficient to make up for that loss will depend upon what the individual reader wishes to get from the book.

University of Illinois LEIGH M. TRIANDIS

Leadership Dynamics and the Trade-Union Leader. By LOIS MAC DONALD with the assistance of MURRAY B. NESBITT, PETER F. FREUND and SAMUEL N. SEIDEMAN. New York: New York University Press, 1959. Pp. xi, 152. $2.50.

This volume is a review of the social-psychological and sociological literature related to leadership with specific emphasis on studies of union leadership. It ends with a brief chapter which attempts an evaluation of the current trends.

The usefulness of this book is very limited. It seems to the present writer that a successful review must either organize many studies in a meaningful whole or discuss a number of important studies in sufficient detail to give a clear picture of the various independent attempts to understand a problem-area. This book has not succeeded by either criterion. It lacks any sort of systematic point of view and its treatment is too brief. Neither the student not the specialist will find this book useful. The student will desire more detail and the specialist will not find much that is new. Also, the organization of the chapters is poor. For instance, the chapter on leadership theory has seven non-communicating, unintegrated sections entitled "The Social Psychologists," "The Human Relations School," "The Emergence of Group Dynamics," "The Development of Sociometry," "Small Group Research," "'Leaderless Discussion Groups," and "The Ohio State Leadership Studies." The logic of this organization escapes this reviewer. On a number of occasions the authors appear naïve in their treatment of social psychological topics. For instance, we are told that sociometry is "a relatively underdeveloped technique for measuring interpersonal relations by recording attitudes and responses" (p. 12) and no reference is made to Festinger's use of matrix algebra and Harary's use of a graph theory with sociometric data. No wonder it is "underdeveloped"! There are important omissions from the literature reviewed. For instance, there is no reference to the important work of Fiedler on interpersonal perceptions and leadership or to the work of the Rosens on the attitudes of union members and leaders. Furthermore, the book is studded with undefined terms, which a student will find very unenlightening. For instance, we are told that Lewin "makes use of topological-spatial and dynamic-vector concepts" which are nowhere defined. Some sentences are in-

comprehensible in the context in which they are presented, though they probably made sense in the original article from which the authors quoted them. For instance, on page 48 we read "Thus, in accordance with the techniques of factor analysis so essential to the *LGD* (Leaderless Group Discussion) method, where changes in the situations are reduced to a minimum from test to retest, consistency correlations and reliability coefficients increase considerably." The sentence is interjected between sentences that are unrelated to it. As it stands it conveys no information.

In all fairness to the authors, however, it must be stated that their discussions of the changing roles of union leaders, and their descriptions of union leader behavior are excellent. It is only when they discuss technical problems and psychology that their effectiveness is greatly diminished.

University of Illinois HARRY C. TRIANDIS

The Onset of Stuttering: Research Findings and Implications. By WENDELL JOHNSON and ASSOCIATES. Minneapolis, University of Minnesota Press, 1959. Pp. ix, 276. Appended summary tables of data, 234 pp. $5.00.

In this book the foremost scientific investigator of stuttering of our time summarizes the results of three interview and case study investigations of stuttering children and their parents conducted by him and his co-workers at the University of Iowa from 1934 to 1957. The most important findings of these studies, based on a total of 246 stuttering subjects and 246 control cases, may be summarized as follows: (1) The two groups of children were essentially similar with respect to substantially all of a multitude of physical and developmental attributes studied. (2) The parents of the stutterers revealed some tendency to impose higher standards of fluency on their children, to be more demanding with respect to weaning, toilet training, and other aspects of child rearing and development, and to be generally more perfectionistic, dissatisfied, and striving than the parents in the control group. (3) The earliest stutterings as they were described by the parents tended to be 'impressively similar' to the normal nonfluencies reported by the parents of most of the non-stutterers.

A focus is provided for these observations by the author's theory, first advanced in 1942, that stuttering is caused largely by the mis-evaluation of normal childhood nonfluency as a speech disorder by overanxious or perfectionistic parents. In this volume Johnson's theory is re-formulated as the concept that shuttering is most appropriately regarded, not as a feature of a person's speech, but as a perceptual and evaluative problem that arises through a process of interaction of a listener and a speaker. The basic assumption in either case is that when the child is first regarded by his parent as a stutterer his speech does not usually differ in important respects from the speech of children who do not come to be regarded as stutterers. This is obviously a difficult assumption to check by direct observation. Johnson's studies represent, in part, an attempt to test it through the reports and descriptions of parents some time after the onset of the problem. Those who are inclined to dispute this will point out that the observation that much of the language used by two groups of parents to describe the speech behavior of their children is the same or similar is very far from equivalent to the observation that the behavior itself was the same.

Facts are facts, however, and, if there are abundant grounds for regarding the

author's interpretations as controversial, there can be little doubt that the mountain of data on the medical and developmental histories of stutterers and the attitudes and adjustments of their parents which he and his co-workers have collected represents one of the most notable research achievements in the area of speech disorders since the beginning of scientific interest in such problems.

Brooklyn College OLIVER BLOODSTEIN

Physical Disability: A Psychological Approach. By BEATRICE A. WRIGHT. New York, Harper and Brothers, 1960. Pp. xx, 408. $6.00.

Physical Disability has many missions to accomplish. It is designed first to present a critical review and evaluation of the research, theory, and empirical findings for the professional person interested in the social-psychological aspects of physical disability. It is, I believe, hoped that those who have ignored this potentially major repository of knowledge about psychological forces, motives, and attitudes revealed by study of the disabled will be attracted by the appeal and challenges posed by this book. It is clear that relatively few psychologists have chosen as yet to work in these fruitful areas, whether in research or applied aspects. The implications of this fact are, until now at least, inadequately explored, but it is suggested that it is not the scarcity of interesting problems that is the major reason for psychologists ignoring the disabled. Dr. Wright's discussion of the disabled person's inferior status position, the values placed on physical wholeness in our society, and community attitudes toward the disabled may all be significant in understanding the retarded development of research and application. Dr. Wright does indicate that the number of workers is slowly increasing, and this is evident in the recent development of rehabilitation centers.

The professional worker in rehabilitation, whether he be physician, occupational therapist, social worker, prosthetist or psychologist, should read the contributions of general psychology and social psychology with equal interest and reward. Frustration-theory, the development of the self concept and the perception of interpersonal relations have theoretical significance that will help explain the behavior and performance of the handicapped client. Rather less rewarding are the case illustrations with which the book is replete. The cases are drawn almost exclusively from the writings of disabled persons about themselves and their experiences. The case material is rich, but herein also lie its weaknesses. It does not, to this reviewer at least, appear to be a very representative sample of the disabled population. The case material in the book might well have been supplemented by case material seen through an observer's eyes.

The mission of *Physical Disability* for the lay person, whether he be disabled, the relative of a disabled person, or any lay member of the community, is especially well accomplished. The meaning and importance of physique are placed in more appropriate perspective. Atypicality is discussed broadly and wisely, and concepts of normality and adjustment are reëxamined.

Dr. Wright's special feeling for the subject of her book and the people in it reveals constantly her dedication to expanding the frontiers of psychological knowledge for an important segment of our society.

Institute for the Crippled and Disabled, New York, N.Y. HAROLD CHENVEN

De l'actinie à l'homme: II. De l'instinct animal au psychisme humain. By HENRI PIÉRON. Paris, Presses Universitaires de France, 1959. Pp. 264. N.F. 14.

This is the second and final volume of Professor Piéron's collected studies on comparative psychophysiology. It constitutes a systematic and comprehensive essay in comparative psychology. It is also a document of great historical interest. Although the mere reprinting of some of these widely scattered articles would in itself prove very valuable to specialists, this new arrangement is more than the simple accumulation of prior publications: it is a living demonstration of the well known principle that the whole is more than the total sum of its parts.

The book is well organized and the short introductory notes to each one of its four parts and to some of its chapters emphasize the comparative and evolutionary aspects of animal behavior and substantiate the title. In recent years, there has been a tendency among psychologists to disregard the really comparative approach for the study of a few preferred species. Piéron's well-informed and meticulous study of multivaried forms of animal behavior should serve as an excellent reminder of the great advantages of the comparative method.

From a historical point of view, some of these classical papers bring us back to the first years of experimental psychology. It must be said, however, that most of these experimental studies retain their usefulness: the facts have not changed and several of these observations have not been repeated. In a few instances, the discussion centers around problems which have had a decisive importance in the development of psychology. The dates of these articles as well as the tone of discussion make us wonder whether Piéron was not a behaviorist even before Watson.

To give an idea of the variety of the information gathered in these pages, let us mention but a few of the problems discussed: the chemical senses and their evolution in several species; geotropism; death-feigning; autotomy and self-mutilation; sensory affectivity; social factors in human development; conditioning and personality. Even though some of these early contributions deal with the notion of instinct, the James-Lange doctrine, Cannon's theory of emotion, and several disputed questions, they are not obsolete in any sense; Piéron's latest contributions, for instance that on the neurophysiology of pain, take into account the great progress made in our neurophysiological knowledge during the last twenty years and prove that the facts of yesterday were really a prelude to the theory of tomorrow.

Université de Montréal

DAVID BÉLANGER

Introduction to Statistical Inference: A Non-Mathematical Exposition of the Theory of Statistics Written for Experimental Scientists. By JEROME C. R. LI. Ann Arbor, Edwards Brothers, 1957. Pp. xiii, 553, $7.50.

Professor Li's text is based on the premise that experimental scientists have neither an adequate mathematical background for the acquisition of statistical material, nor the formal educational time to acquire such a background. A set of theorems is presented which are not mathematically proven but rather empirically verified by the use of sampling procedures. For example, 1000 samples of size 5 are drawn from a normal population of integers with known mean and variance. For each sample, the sum of squared deviations about the population mean is computed and divided by the population-variance. The distribution of these quantities is reported and compared with the appropriate Chi-square distribution, verifying that the

quantity is distributed as Chi-square. With this approach as its keystone, the book covers a variety of topics, from means and variances to randomized block designs.

The text is readable and extensively illustrated. The chapters on Chi-square, *t*-, and *F*-distributions are well developed, and should give the student an understanding of the relationships among these distributions. Unfortunately, certain important topics receive little or no attention, while others receive too much consideration. There is no development of probability theory, the foundation of statistical inference. There is an extensive discussion of the property of unbiasedness of estimators, but no mention of the more important properties of consistency and relative efficiency. The linear model for analysis of variance is extensively treated but only bare mention is made of the equally important random and mixed models. There are five pages on the Least Significant Difference test, a partly unnecessary presentation in light of the subsequent discussion of the Duncan Range-Test.

The reviewer takes exception to the premise that experimental scientists (or at least experimental psychologists) do not and can not attain a level of mathematical competence sufficient to grapple with, for example, the usual mathematical statistics course. Nor does the reviewer agree with the author's belief that the book will prepare the student for the later acquisition of specialized material. In view of the weaknesses cited above (probability theory, properties of estimators, discussion of random and mixed models), it is difficult to credit this.

University of Massachusetts — JEROME L. MYERS

Cybernetics and Management. By STAFFORD BEER. New York, John Wiley and Sons, Inc., 1959. Pp. xviii, 214. $4.50.

The intelligent layman who wishes to read an introduction to cybernetics and a discussion of its applications to management problems will find this small volume useful. The author is Head of the Department of Operational Research and Cybernetics of United Steel Companies, Ltd., in England. The book grew out of his lectures to the Royal Swedish Academy of Engineering Sciences and the International Association for Cybernetics. It presents the basic notions of cybernetics, discusses the theoretical possibilities of this field of knowledge, and describes a number of possible applications to management problems.

The book is clearly written and requires very little mathematical sophistication. Those who found Wiener's book somewhat difficult may welcome the publication of Beer's work. On the other hand, those who know the current work in this field and are aware of the contributions of Ashby, Shannon, Simon, and Von Neuman, will find little that is new to them in the present account.

Psychologists will find certain chapters particularly interesting. A discussion of the theory of automata, and chapters on adaptive behavior, on teaching machines, and on "amplification of intelligence" by means of suitable man-computer systems are likely to intrigue many psychologists. These chapters include references to the work of Bush and Mosteller, and some work on teaching machines by Pask which seems highly promising.

Industrial psychologists may find the last part of the book of some interest. The author argues that management means control; since cybernetics is basically the study of the principles of controlling complicated systems, some managers will find these principles relevant to their work.

University of Illinois — H. C. TRIANDIS

The Measurement of Values. By LEWIS LEON THURSTONE. Chicago, The University of Chicago Press, 1959. Pp. vii, 322. $7.50.

This volume represents a compilation of Thurstone's important published contributions to the problems of psychological scaling and attitude measurement. He himself selected most of the articles to be included, and wrote summary statements to introduce each section. After his death, Mrs. Thurstone completed the task of editing. Three additional proposed sections on learning, factor analysis, and testing theory had to be omitted, and editing for redundancy and typographical errors was limited. The collection of these sometimes hard-to-find original articles is, nevertheless, both useful and informative.

Thurstone expressed clearly his intentions to help "to introduce quantitative rational theory into psychology." These 27 papers furnish a basis for evaluation of the progress he made in these areas toward his goal.

The creativity of his single contributions is much more evident in his own papers than in reports of them in secondary sources which do not take into account the period of their publication. This is particularly true for the 1927 article on "Psychophysical Analysis," in which he explored the discriminal dispersion and its implications, with clarity and insight. This article was, in his judgment, his "best contribution to psychology." This reviewer agrees.

The research worker who accepts the quantitative goal will find it profitable to read this sequence of papers. He will not only be encouraged by the real progress shown, but also be stimulated toward further research. The changes in direction shown in the later papers acknowledge the importance of the prediction of a criterion of actual behavior in an everyday situation, and of the establishment of the laws by which metrics can be combined to make such predictions.

Cornell University PATRICIA C. SMITH

Words and Object, by WILLARD VAN ORMAN QUINE. New York, Published jointly by The Technology Press of the Massachusetts Institute of Technology and John Wiley and Sons, Inc. Pp. xv, 294, $5.50.

Language in Relation to a Unified Theory of the Structure of Human Behavior, Part III (chapters 11-17). By KENNETH L. PIKE. Preliminary edition issued by the Summer Institute of Linguistics, Glendale, California. Pp. viii, 146. $3.50.

These are highly technical books which the psychologist will find baffling unless he is familiar with recent developments in logic and in linguistics. Each of these disciplines has been evolving a highly specialized language which must be patiently learned before one can understand what is being said, and neither Quine nor Pike has made the slightest concession to the uninitiated reader. Perhaps this is necessary for efficient communication on progressively higher levels of abstraction. Certainly this is true of mathematics; and certainly the psychologist, of all people, has no right to condemn another discipline for having developed its own peculiar jargon. The reviewer reports with pride that, after considerable effort, he was able to make sense out of both books and that he found the reward worth the effort.

The encouraging thing for the psychologist is to find the interests of an eminent mathematical logician and an eminent linquist converging on what ought to be one of psychology's central problems; namely, the nature of human communication. The modest contributions of the psychologists are dealt with by both authors with charity, and by Pike with great erudition, and these are related meaningfully to the

technical problems of the two disciplines. The reviewer's judgment is that the psychological student of language can still learn much from modern logic and linguistics.

Cornell University — R. B. MacLeod

Psychology in Nursing. By Wendell Cruze. 2nd ed., New York, McGraw-Hill Book Company, 1960. Pp. vi, 536. $6.00.

This revision of a text in wide use in schools of nursing is written for the beginning student with little or no prior contact with the field of psychology. The traditional subjects of a survey course are treated without involving the student in complex theories, conecpts, or techniques. It contains an admixture of material in anatomy and physiology; personality-growth, adaptation and deviation; cognitive functions; as well as advice and admonitions of a wise teacher to his youthful charges entering the nursing profession.

The book's merits and deficiencies are inter-related. Detailed information is not pursued intensively. The student will be apprised of many matters but they may slip by easily since emphasis is not strong. At the same time, because the book is apparently directed to high school graduates, it does not pretend to match strides with other standard works for college courses in general psychology.

The friend-at-your-side manner of the author is a splendid shoehorn into psychology. He does not, however, offer sufficient cautions to the reader that general principles presented here may be contradicted by experience with individual cases, and that variability of response is as much expected as are uniformities. The material on aging is considerably oversimplified even for the young nursing student. The new chapter on "Adjustment Problems of the Student Nurse" should be helpful, if only to guide the student to a counselor or to other reading. The bibliography and references have been brought up to date in several chapters but only 17 of the 132 references are post-1955.

Albert Einstein College of Medicine, New York City — Bernard Kutner

Psychological Problems in Mental Deficiency. By Seymour B. Sarason. Third edition. New York, Harper and Brothers, 1959. Pp. 652. $6.50.

This book is a reprinting of two of Sarason's books, which have been combined, without revision, into a single book. Apparently the same galleys were used from the second edition of *Psychological Problems in Mental Deficiency,* for page by page the second and third editions are duplicates. What makes this a "third edition" is that following the unchanged second edition of Sarason's book there is added the entire monograph by Sarason and Thomas Gladwin, also previously published and here unchanged: *Psychological and Cultural Factors in Mental Subnormality.*

Perhaps Sarason himself has made the best appraisal of this "third edition" when he states in the preface that "it would have been desirable if, in this third edition, the monograph and book could really have been integrated."

Since both publications comprising this third edition have been reviewed previously,[1] there is little need to add more here except to remind the reader that these are good scholarly works and constitute a comprehensive review of the research literature. There is now the added convenience of having both publications bound into one book. A fourth edition which could rival the *Handbook of Experimental Psy-*

[1] See this Journal, 63, 1950, 313-314.

chology in bulk would be to append to the present third edition Masland, Sarason, and Gladwin's bulky *Mental Subnormality.*

The University of Texas AUSTIN E. GRIGG

Thought and Action: A Physiological Approach. By RICHARD K. OVERTON. New York, Random House, 1959. Pp. 116. $0.95.

This essay in speculative neurologizing uses the reverberating circuit to "unite the field of psychology by giving it a biological foundation" (p. 5). Reverberating circuits, it is argued, explain thought, memory, insight, drives, attention, and the behavior disorders. No independent evidence is given for their existence except the behavioral ones and the argument reduces to the tautology: behavior results from reverberating circuits, and reverberating circuits are inferred from behavior. Beyond this the book has serious faults. The physiology used here is more metaphorical than real. More important in a book supposedly intended for undergraduates is the debatable use of psychological information. Controversial issues are given gratuitous explanations; experiments not even cited in the references are given arguable conclusions; the author's published and unpublished studies are presented as fact when the conclusions drawn from them are at variance with considerable information which already exists; and 'reasons' are offered to explain everyday events which have nothing in common with the experimental evidence used to explain them. The sad thing is that the author, having set out to write a stimulating book, has succeeded only in compounding a hodge-podge of physiology and psychology such that the student who masters this book will find himself faced with a considerable job of future unlearning.

Cornell University ALLAN C. GOLDSTEIN

Einführung in die Pharmako-psychologie. By HERBERT LIPPERT. Hans Hubert, Bern, 1959. Pp. 254. DM 32.

This is an excellent book, encyclopedic in nature. It is recommended to any one seeking information about the present status and the history of knowledge concerning the effects of drugs. The textual matter, consisting of 180 pages is followed by 70 pages containing references to the relevant literature (1695 titles) and a very comprehensive index.

The text is divided into 12 chapters as follows: (a) Introduction, which credits Kraepelin as being the pioneer of this young science. (b) Motor and sensory effects distinguished. (c) Abolishing pain and arousing ecstasy. (d) Perception, hallucination, space, time, and memory. (e) Thinking and willing. (f) Activity in useful work. (g) Expression of the personality. (h) Clear consciousness and the dream consciousness during sleep. (i) Characterology, racial and geographical differences, *e.g.* hemp grown in cool Europe produces no drug. (j) National differences, *e.g.* the single intoxicating drug in China is opium smoke. (k) Reference to age from childhood to senility. (l) Religion, mythology, fine arts, politics, and surgical use. (m) Methodology, drugs and placebos in animal experiments. (n) The comparative chemical structure of drugs.

The usefulness of the book is far from being restricted to the scientific specialist. The text is highly interesting also to the the general reader.

Miami, Florida MAX F. MEYER

The History and Origin of Language. By A. S. Diamond. New York, Philosophical Library, 1959. Pp. 280, $7.50.

Psychologists have more to read than they have time for, and will therefore be glad to know that they can afford to ignore this book. It is unsound in almost every detail of fact, almost every direction of interpretation, almost every line of inference. About fifty years ago, the Linguistic Society of Paris established a standing rule barring papers on the origin of language from its sessions. This ruling—undemocratic and deplorable as we might think it—was directed squarely against such time-wasting speculation as is found in the volume under review.

This is not to say that we may not, in time, be able to discover something of the origin of language. Such discovery cannot, however, come merely from the examination of language itself (particularly when language itself is viewed inaccurately, as it is by the present author). It requires a frame of reference large enough to compare, with considerable accuracy, human language and the communicative systems of other animals. The interested reader should consult the reviewer's article "Logical Considerations in the Study of Animal Communication," in *Animal Sounds and Communication,* American Institute of Biological Sciences, 1960.

Cornell University — Charles F. Hockett

Methods of Correlation and Regression Analysis: Linear and Curvilinear. By Mordecai Ezekiel and Karl A. Fox. 3rd ed. New York, John Wiley and Sons, Inc., 1959. Pp. xv, 548, $10.95.

Although the title is slightly changed and Ezekiel has been joined by Fox, the book is readily, perhaps too readily, recognized as the third edition of the classic, *Methods of Corelation Analysis,* originally published in 1930 and revised in 1941. The new chapters on analysis of variance in relation to regression and the solution of simultaneous equations are useful but not distinguished. The same may be said of the extensively revised chapters on standard errors in multiple regression and on standard error formulas for time-series. Although the book has merit, too many excellent texts have appeared in the 29-year interval since the first edition of this book, and the changes in this edition have not been sufficient for it to retain its original preëminence.

The University of Wisconsin — David A. Grant

Modern Probability Theory and Its Applications. By Emanuel Parzen. New York, John Wiley and Sons, Inc., 1960. Pp. xv, 464. $10.75.

This is one of the most suitable books for use as a text in a first course in probability-theory. While its aim is not the detail or depth to be found in Feller's classical text, the latter is limited in its first volume (which is all that has appeared) to considerations in the discrete case of probability, while Parzen's book considers the continuous case, too. The author's writing is clear, and the novice will gain insight from the style of exposition. There are good examples. The level of sophistication varies from the simplest concepts in the earlier chapters to proofs of standard limit theorems near the end of the book. Besides being useful as a text in a mathematics course, the book will undoubtedly serve as a useful reference on the basic ideas of probability theory for the psychologist with sufficient mathematical training (calculus).

Cornell University — J. Kiefer

Road of Propaganda: The Semantics of Biased Communication. By KARIN DOVRING. Introduction by Harold D. Lasswell. New York, Philosophical Library, 1959. Pp. 158. $4.75

This book consists of a series of numbered (37) thoughtful observations by a noted Swedish scholar about the nature and effectiveness of biased communications in international settings. It presents no new research, nor does it attempt a serious theoretical integration of existing research. Although considerable theory and research in 'mass-communications psychology' is referred to, strictly speaking this should not be classed as a volume in psychology. Nevertheless, students of social psychology or mass communications may find some observations of interest and some ideas for research in the volume.

Stanford University NATHAN MACCOBY

INDEX

CHARLES VALLEY BROOK, University of Texas

AUTHORS

(The names of authors of original articles are printed in CAPITALS; of authors of books reviewed, in roman; and of reviewers, in *italics.*)

SUBJECTS

(References in *italic* figures are to reviews.)

PRICE LIST OF THE BACK FILES
THE AMERICAN JOURNAL OF PSYCHOLOGY

Volume	Year	Price per number 1	2	3	4	Price per volume
I	1887–88	$2.00	$2.00	$—	$5.00	$ —
II	1888–89	2.00	2.00	2.00	5.00	11.00
III	1890–91	(4.00)	4.00	4.00	—	—
IV	1891–92	4.00	4.00	4.00	—	—
VII	1895–96	1.85	1.85	1.85	1.85	7.00
VIII	1896–97	1.85	1.85	1.85	1.85	7.00
IX	1897–98	1.85	1.85	1.85	1.85	7.00
X	1898–99	1.85	1.85	1.85	1.85	7.00
XI	1899–00	1.85	1.85	1.85	1.85	7.00
XII	1900–01	1.85	1.85	1.85	1.85	7.00
XIII	1902	1.85	1.85	1.85	1.85	7.00
XIV	1903	1.85	1.85	1.85	1.85	7.00
XV	1904	1.85	1.85	1.85	1.85	7.00
XVI	1905	1.85	—	—	1.85	—
XVII	1906	1.85	1.85	1.85	1.85	7.00
XVIII	1907	1.85	1.85	1.85	1.85	7.00
XIX	1908	1.85	1.85	1.85	1.85	7.00
XX	1909	1.85	1.85	—	1.85	—
XXI	1910	2.00	(2.00)	—	1.85	—
XXII	1911	—	1.85	—	—	—
XXIII	1912	(2.00)	1.85	(2.00)	(2.00)	7.00
XXIV	1913	—	(2.00)	—	(2.00)	—
XXV	1914	(2.00)	1.85	—	2.00	—
XXVI	1915	1.85	—	—	1.85	—
XXVII	1916	(2.00)	1.85	—	—	—
XXVIII	1917	(2.00)	(2.00)	2.00	(2.00)	7.50
XXIX	1918	(2.00)	(2.00)	2.00	2.00	7.50
XXX	1919	—	—	(2.00)	(2.00)	—
XXXI	1920	—	(2.00)	(2.00)	—	—
XXXII	1921	—	(2.00)	(2.00)	—	—
XXXIII	1922	—	(2.00)	(2.00)	(2.00)	—
XXXIV	1923	1.85	1.85	1.85	1.85	7.00
XXXV	1924	1.85	1.85	1.85	1.85	7.00
XXXVI	1925	1.85	1.85	1.85	1.85	7.00
XXXVII	1926	1.85	1.85	1.85	1.85	7.00
XXXVIII	1927	1.85	1.85	1.85	1.85	7.00
XXXIX	1927	Washburn Commemorative Volume				5.50
XL	1928	1.85	1.85	1.85	1.85	7.00
XLI	1929	1.85	1.85	1.85	1.85	7.00
XLII	1930	1.85	1.85	1.85	1.85	7.00
XLIII	1931	1.85	1.85	1.85	1.85	7.00
XLIV	1932	1.85	1.85	1.85	1.85	7.00
XLV	1933	1.85	1.85	1.85	1.85	7.00
XLVI	1934	1.85	1.85	1.85	1.85	7.00
XLVII	1935	1.85	1.85	1.85	1.85	7.00
XLVIII	1936	1.85	1.85	—	1.85	—
XLIX	1937	1.85	1.85	1.85	1.85	7.00
L	1937	Golden Jubilee Volume				5.50
LI	1938	1.85	1.85	1.85	1.85	7.00
LII	1939	1.85	1.85	1.85	1.85	7.00
LIII	1940	1.85	1.85	1.85	1.85	7.00
LIV	1941	1.85	1.85	1.85	1.85	7.00
LV	1942	1.85	1.85	1.85	1.85	7.00
LVI	1943	1.85	1.85	1.85	1.85	7.00
LVII	1944	1.85	1.85	1.85	1.85	7.00
LIX	1946	—	1.85	1.85	1.85	—
LX	1947	1.85	1.85	1.85	1.85	7.00
LXI	1948	1.85	1.85	1.85	1.85	7.00
LXII	1949	1.85	1.85	1.85	1.85	7.00
LXIII	1950	1.85	1.85	1.85	1.85	7.00
LXIV	1951	1.85	1.85	1.85	1.85	7.00
LXV	1952	1.85	1.85	1.85	1.85	7.00

Parentheses designate reprinted numbers. The reprinting is done by the anastatic process, consequently the reproduction is made with photographic accuracy.